GenCore version 4.5

OM nucleic nucleic search, using sw model

Run on: July 2, 2002, 19:28:24 ; Search time 1898.13 Seconds
(without alignments)
661.489 Million cell updates/sec

Title: US-09-475-704-1
Perfect score: 60
Sequence: 1 gacatcaaagcaggaggcccaa.....tggaccgcttcttccaggacc 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 1046268293 residues

Total number of hits satisfying chosen parameters: 4595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: qb_ba:*
2: qb_htg:*
3: qb_in:*
4: qb_om:*
5: qb_ov:*
6: qb_pat:*
7: qb_ph:*
8: qb_pl:*
9: qb_pr:*
10: qb_ro:*
11: qb_sts:*
12: qb_sy:*
13: qb_un:*
14: qb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	55.2	92.0	1608	6	AX188558	AX188558 Sequence
2	55.2	92.0	1914	6	AX188560	AX188560 Sequence
3	55.2	92.0	2493	6	AX188562	AX188562 Sequence
4	55.2	92.0	4450	6	AX188564	AX188564 Sequence
5	53.6	89.3	1509	12	AF201927	AF201927 Synthetic
6	53.6	89.3	1847	12	AF202464	AF202464 Synthetic
7	53.6	89.3	1847	12	AF202465	AF202465 Synthetic
8	52	86.7	4288	6	AX149648	AX149648 Sequence
9	50.4	84.0	1548	12	AF287354	AF287354 Synthetic
10	50.4	84.0	4352	12	AF287352	AF287352 Synthetic
11	50.4	84.0	4353	12	AF287353	AF287353 Synthetic
12	47.2	78.7	1539	6	AX306431	AX306431 Sequence
13	47.2	78.7	3981	6	AX306430	AX306430 Sequence
14	47.2	78.7	4441	6	AX306429	AX306429 Sequence
15	47.2	78.7	4443	6	AX306428	AX306428 Sequence
16	43	71.7	4307	6	AX019132	AX019132 Sequence
17	43	71.7	4307	6	AX035453	AX035453 Sequence
18	43	71.7	4307	6	AX056836	AX056836 Sequence
19	43	71.7	4307	6	AX283597	AX283597 Sequence
20	43	71.7	4307	6	AX283607	AX283607 Sequence
21	43	71.7	4327	6	AX035465	AX035465 Sequence
22	43	71.7	4353	6	AX035464	AX035464 Sequence
23	43	71.7	4642	6	AX035463	AX035463 Sequence
24	43	71.7	9772	6	AX283584	AX283584 Sequence
25	39.2	65.3	669	14	HIB274572	AJ274572 Human imm
26	38.2	63.7	393	14	AF196450	AF196450 HIV 1 iso
27	38.2	63.7	424	14	AF184568	AF184568 HIV 1 iso
28	38.2	63.7	690	14	AF184569	AF184569 HIV 1 iso
29	38.2	63.7	693	14	AB034323	AB034323 Human imm
30	38.2	63.7	769	14	AF184519	AF184519 HIV 1 iso
31	38.2	63.7	1491	14	AF391263	AF391263 HIV 1 clo
32	38.2	63.7	8855	14	AF361879	AF361879 HIV 1 iso
33	38.2	63.7	9002	14	AF443086	AF443086 HIV 1 iso
34	38.2	63.7	9009	14	AF286224	AF286224 HIV-1 str
35	38.2	63.7	9010	14	AF110974	AF110974 HIV-1 iso
36	38.2	63.7	9010	14	AF110975	AF110975 HIV-1 iso
37	38.2	63.7	9012	14	AF110973	AF110973 HIV 1 iso
38	38.2	63.7	9064	14	AF443094	AF443094 HIV 1 iso
39	38.2	63.7	9788	14	AF103818	AF103818 Simian im
40	37.6	62.7	349	14	HIV228276	AJ228276 Human imm
41	37.6	62.7	393	14	AF106388	AF106388 HIV 1 iso
42	37.6	62.7	616	14	AF184555	AF184555 HIV 1 iso
43	37.6	62.7	616	14	AF184560	AF184560 HIV 1 iso
44	37.6	62.7	617	14	AF184570	AF184570 HIV 1 iso
45	37.6	62.7	625	14	AF184583	AF184583 HIV 1 iso

ALIGNMENTS

RESULT 1
AX188558
LOCUS AX188558 1608 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 27 from Patent WO0147955.
ACCESSION AX188558
VERSION AX188558.1 GI:15142199
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 1608)
AUTHORS Hanke,T.M. and Mcmichael,A.J.
TITLE Improvements in or relating to immune responses to hiv
JOURNAL Patent: WO 0147955-A 27 05-JUL-2001;
MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine Initiative (US) ; University of Nairobi (KE)
FEATURES Location/Qualifiers
source 1..1608
/organism "synthetic construct"
/db_xref "taxon:32630"
/note "Chimeric polynucleotide"
BASE COUNT 314 a 598 c 458 g 238 t
ORIGIN

Query Match 92.0%; Score 55.2; DB 6; Length 1608;
Best Local Similarity 95.0%; Pred. No. 7.4e-06;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gacatcaagcagggccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 472 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 531

RESULT 2
AX188560
LOCUS AX188560 1914 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 29 from Patent WO0147955
ACCESSION AX188560
VERSION AX188560.1 GI:15142200
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequence
REFERENCE 1 (bases 1 to 1914)
AUTHORS Hanke,T.M. and Mcmichael,A.J.
TITLE Improvements in or relating to immune responses to hiv
JOURNAL Patent: WO 0147955-A 29 05 JUL 2001;
MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
Initiative (US) ; University of Nairobi (KE)
FEATURES Location/Qualifiers
source 1. .1914
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Chimeric polynucleotide"
BASE COUNT 383 a 728 c 534 g 269 t
ORIGIN

Query Match 92.0%; Score 55.2; DB 6; Length 1914;
Best Local Similarity 95.0%; Pred. No. 7.3e-06;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gacatcaagcagggccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 466 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

RESULT 3
AX188562
LOCUS AX188562 2493 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 31 from Patent WO0147955.
ACCESSION AX188562
VERSION AX188562.1 GI:15142201
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 2493)
AUTHORS Hanke,T.M. and Mcmichael,A.J.
TITLE Improvements in or relating to immune responses to hiv
JOURNAL Patent: WO 0147955-A 31 05-JUL-2001;
MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
Initiative (US) ; University of Nairobi (KE)
FEATURES Location/Qualifiers
source 1. .2493
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Chimeric polynucleotide"
BASE COUNT 516 a 887 c 685 g 405 t
ORIGIN

Query Match 92.0%; Score 55.2; DB 6; Length 2493;
Best Local Similarity 95.0%; Pred. No. 7e-06;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gacatcaagcagggccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 466 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

RESULT 4
AX188564
LOCUS AX188564 4350 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 33 from Patent WO0147955.
ACCESSION AX188564
VERSION AX188564.1 GI:15142202
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 4350)
AUTHORS Hanke,T.M. and Mcmichael,A.J.
TITLE Improvements in or relating to immune responses to hiv
JOURNAL Patent: WO 0147955-A 33 05 JUL 2001;
MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
Initiative (US) ; University of Nairobi (KE)
FEATURES Location/Qualifiers
source 1 4350
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Chimeric polynucleotide"
BASE COUNT 888 a 1630 c 1259 g 573 t
ORIGIN

Query Match 92.0%; Score 55.2; DB 6; Length 4350;
Best Local Similarity 95.0%; Pred. No. 6.6e-06;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gacatcaagcagggccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 466 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

RESULT 5
AF201927
LOCUS AF201927 1509 bp DNA linear SYN 16-MAR-2000
DEFINITION Synthetic construct gag protein gene, complete cds.
ACCESSION AF201927
VERSION AF201927.1 GI:7248702
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 1509)
AUTHORS zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
Selby,M., Otten,G.R. and Barnett,S.W.
TITLE Increased expression and immunogenicity of sequence-modified human
immunodeficiency virus type 1 gag gene
JOURNAL J. Virol. 74 (6), 2628-2635 (2000)
MEDLINE 20148954
PUBMED 10684277
REFERENCE 2 (bases 1 to 1509)
AUTHORS zur Megede,J. and Barnett,S.W.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,
Emeryville, CA 94608, USA
FEATURES Location/Qualifiers
source 1. .1509
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS 1. .1509
/note="HIV-1SF2 p55gag"
/codon_start=1
/transl_table=11
/product="gag protein"
/protein_id="AAF43628.1"

```
                     /db_xref "GI:7248703"
                     /translation="[illegible]
                     AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
                     [illegible]
                     VEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVH
                     PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
                     RMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKT
                     ILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNPATIMMQRGNFRNQRKIV
                     [illegible]
                     LQSRPEPTAPPEESFRFGEEKTTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
BASE COUNT      328 a    543 c    479 g    159 t
ORIGIN

  Query Match             89.3%;  Score 53.6;  DB 12;  Length 1509;
  Best Local Similarity   93.3%;  Pred. No. 2.1e-05;
  Matches   56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||||   |||||||||||||||||||||||||||||||||||||||||| |||||||
Db        856 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 915


RESULT 6
AF202464
LOCUS       AF202464                1847 bp    DNA     linear   SYN 12-MAR-2000
DEFINITION  Synthetic construct gag protease fusion protein GP1 gene, complete
            cds.
ACCESSION   AF202464
VERSION     AF202464.1  GI:7229424
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 1847)
  AUTHORS   zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
            Selby,M., Otten,G.R. and Barnett,S.W.
  TITLE     Increased expression and immunogenicity of sequence-modified human
            immunodeficiency virus type 1 gag gene
  JOURNAL   J. Virol. 74 (6), 2628-2635 (2000)
  MEDLINE   20148954
   PUBMED   10684277
REFERENCE   2  (bases 1 to 1847)
  AUTHORS   zur Megede,J. and Barnett,S.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton
            Street, Emeryville, CA 94608, USA
FEATURES             Location/Qualifiers
     source          1..1847
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             1..1509
                     /codon_start=1
                     /transl_table=11
                     /product="gag protease fusion protein GP1"
                     /protein_id="AAF42819.1"
                     /db_xref="GI:7229425"
                     /translation="MGARASVLSGGELDKWEKIRLRPGGKKKYKLKHIVWASRELERF
                     AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
                     [illegible]
                     VEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVH
                     PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
                     RMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKT
                     ILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNPATIMMQRGNFRNQRKIV
                     [illegible]
                     LQSRPEPTAPPEESFRFGEEKTTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
BASE COUNT      420 a    620 c    579 g    228 t
ORIGIN

  Query Match             89.3%;  Score 53.6;  DB 12;  Length 1847;
  Best Local Similarity   93.3%;  Pred. No. 2.1e-05;
  Matches   56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||||   |||||||||||||||||||||||||||||||||||||||||| |||||||
Db        856 GACATC[illegible]GACTACGTGGACCGCTTCTACAAGACC 915


RESULT 7
AF202465
LOCUS       AF202465                1847 bp    DNA     linear   SYN 13-MAR-2000
DEFINITION  Synthetic construct gag protease fusion protein GP2 gene, complete
            cds.
ACCESSION   AF202465
VERSION     AF202465.1  GI:7229426
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 1847)
  AUTHORS   zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
            Selby,M., Otten,G.R. and Barnett,S.W.
  TITLE     Increased expression and immunogenicity of sequence-modified human
            immunodeficiency virus type 1 gag gene
  JOURNAL   J. Virol. 74 (6), 2628-2635 (2000)
  MEDLINE   20148954
   PUBMED   10684277
REFERENCE   2  (bases 1 to 1847)
  AUTHORS   zur Megede,J. and Barnett,S.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton
            Street, Emeryville, CA 94608, USA
FEATURES             Location/Qualifiers
     source          1..1847
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             1..1509
                     /codon_start=1
                     /transl_table=11
                     /product="gag protease fusion protein GP2"
                     /protein_id="AAF42820.1"
                     /db_xref="GI:7229427"
                     /translation="[illegible]
                     AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
                     [illegible]
                     [illegible]
                     PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
                     [illegible]
                     [illegible]
                     [illegible]
                     LQSRPEPTAPPEESFRFGEEKTTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
BASE COUNT      456 a    576 c    565 g    250 t
ORIGIN

  Query Match             89.3%;  Score 53.6;  DB 12;  Length 1847;
  Best Local Similarity   93.3%;  Pred. No. 2.1e-05;
  Matches   56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||||   |||||||||||||||||||||||||||||||||||||||||| |||||||
Db        856 GACATC[illegible]GACTACGTGGACCGCTTCTACAAGACC 915


RESULT 8
AX149648
LOCUS       AX149648                4238 bp    DNA     linear   PAT 06-JUN-2001
DEFINITION  Sequence 2 from Patent WO0136614.
ACCESSION   AX149648
VERSION     AX149648.1  GI:14348647
KEYWORDS    .
SOURCE      Human immunodeficiency virus.
  ORGANISM  Human immunodeficiency virus
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
```

```
REFERENCE   1  (bases 1 to 4288)
  AUTHORS   Shao,Y., Wagner,R., Wolf,H. and Graf,M.
  TITLE     The genome of the hiv-1 inter-subtype (c/b') and use thereof
  JOURNAL   Patent: WO 0136614-A 2 25-MAY-2001;
            Geneart GMBH Gesellschaft fuer angewandte Boitechnologie (DE) ;
            Shao, Yiming (CN)
FEATURES             Location/Qualifiers
     source          1..4288
                     /organism="Human immunodeficiency virus"
                     /db_xref="taxon:12721"
BASE COUNT     1026 a   1358 c   1397 g    507 t
ORIGIN

  Query Match             86.7%;  Score 52;  DB 6;  Length 4288;
  Best Local Similarity   91.7%;  Pred. No. 5.2e-05;
  Matches   55;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              |||||||||||||||||||||||||||| |||||||| ||| | ||||||||||||
Db        853 GACATCAAGCAGGGCCCCAAGGAGCCCTTCAGGGACTACGTCGACAGGTTCTTCAAGACC 912


RESULT 9
AF287354
LOCUS       AF287354                1548 bp    DNA     linear   SYN 01-NOV-2000
DEFINITION  Synthetic construct HIV-1-derived [illegible] protein gene, complete cds.
ACCESSION   AF287354
VERSION     AF287354.1  GI:11066866
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 1548)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Helper plasmids for production of HIV derived vectors
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1548)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2000) Chemical Pathology, Women's and Children's
            Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES             Location/Qualifiers
     source          1..1548
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="derived from Human immunodeficiency virus type 1
                     isolate YU-2"
     CDS             37..1539
                     /note="codon optimized reading frame"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAG28738.1"
                     /db_xref="GI:[illegible]"
                     /translation="MGARASV[illegible]SAG[illegible]WEKIR[illegible]YR[illegible]KH[illegible]VWAS[illegible]LERF
                     AVDPGLLETSEGCRQI[illegible]QL[illegible]SL[illegible]TGSEELRSLYNTVATLYCVHQ[illegible]IEV[illegible]ALL
                     KIEEEQNKSKKKAQQAAADTGNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVVE
                     EKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPV
                     HAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM
                     YSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNAN[illegible]KTIL
                     KALGPAATLEEMMTACQGVGG[illegible]KARVLAEAMSQVTN[illegible]ATIMMQRGNFRNQRK[illegible]VK[illegible]
                     FNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANF[illegible]GKIWPSHKG[illegible]
                     SRPEPTAPSEESVRFGEETTTPSQKQEPIDKELYPLASL[illegible]"
BASE COUNT     [illegible] a   [illegible] c   [illegible] g   [illegible] t
ORIGIN

  Query Match             84.0%;  Score 50.4;  DB 12;  Length 1548;
  Best Local Similarity   90.0%;  Pred. No. 0.00017;
  Matches   54;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
```

```
              |||||| ||||||||||||||||||||| | ||||||||||| | |||| |||||||
Db        886 GACATCAGGCAGGGCCCCAAGGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 945


RESULT 10
AF287352
LOCUS       AF287352                4352 bp    DNA     linear   SYN 01-NOV-2000
DEFINITION  Synthetic construct HIV-1 derived gag protein and pol protein
            genes, complete cds.
ACCESSION   AF287352
VERSION     AF287352.1  GI:11066861
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4352)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Helper plasmids for production of HIV derived vectors
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4352)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2000) Chemical Pathology, Women's and Children's
            Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES             Location/Qualifiers
     source          1..4352
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="derived from Human immunodeficiency virus type 1
                     isolate YU-2"
     CDS             37..1539
                     /note="codon optimized reading frame"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAG28735.1"
                     /db_xref="GI:11066862"
                     /translation="M[illegible]VL[illegible]AG[illegible]EWEKI[illegible]LRP[illegible]KK[illegible]Y[illegible]KH[illegible]IVWASRELERF
                     AVD[illegible]GLLETSE[illegible]YNTVATLYCVHQKI[illegible]VKDTKEALE
                     K[illegible]NAWVKVVE
                     [illegible]AEWDRLHPV
                     [illegible]NKIVRM
                     Y[illegible]PTSILDIRQGPKEPFRDYVDRFYKT[illegible]RAEQASQEVKNWMTE[illegible]LVQNANPDCKTIL
                     KALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKC
                     [illegible]
                     [illegible]PSQKQEP[illegible]DKELYPLASLRSLFGSDPSSQ"
     CDS             [illegible]..4343
                     /note="codon optimized reading frame"
                     /codon_start=1
                     /transl_table=11
                     /product="pol protein"
                     /protein_id="AAG28736.1"
                     /db_xref="GI:11066863"
                     /translation="[illegible]SEAGA
                     [illegible]MIG
                     [illegible]IKVRQYD[illegible]GHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIET
                     [illegible]
                     [illegible]KKKSVTVLDVGDAYFSVPLHED
                     [illegible]
                     QYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKW
                     TV[illegible]TEVIPLTE
                     [illegible]
                     [illegible]
                     [illegible]
                     [illegible]
                     KVYLAWVPAHKGIGGNEQVDKLVSAGIRKVL[illegible]
                     LPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIE
                     AEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTSATVKAACWWAGIKQEFGIPY
                     [illegible]
                     [illegible]
                     V[illegible]MAGDDCVAGRQDED"
BASE COUNT     1066 a   1344 c   1410 g    532 t
ORIGIN
```

```
  Query Match             84.0%;  Score 50.4;  DB 12;  Length 4353;
  Best Local Similarity   90.0%;  Pred. No. 0.00015;
  Matches   54;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggcccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              |||||| ||||||||||||||||||||||||| |||||||||||| | |||| ||||||
Db        886 GACATCAGGCAGGGGCCCAAGGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 945


RESULT 11
AF287353
LOCUS       AF287353                4353 bp    DNA     linear   SYN 01-NOV-2000
DEFINITION  Synthetic construct HIV-1-derived gag-pol fusion protein gene,
            complete cds.
ACCESSION   AF287353
VERSION     AF287353.1  GI:11066864
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4353)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Helper plasmids for production of HIV derived vectors
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4353)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2000) Chemical Pathology, Women's and Children's
            Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES             Location/Qualifiers
     source          1..4353
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="derived from Human Immunodeficiency Virus type 1
                     isolate YU-2"
     CDS             37..4344
                     /note="codon optimized reading frame"
                     /codon_start=1
                     /transl_table=11
                     /product="gag-pol fusion protein"
                     /protein_id="AAG28737.1"
                     /db_xref="GI:11066865"
                     /translation="M[illegible]VWASRELERF
                     AV[illegible]NVAT[illegible]V[illegible]EVK[illegible]EALE
                     [illegible]
                     EKAFSPEVIPMFSAL[illegible]NTMLNTVGGHQAAMQMLKETINEEAAEWDR[illegible]HPV
                     HAG[illegible]IA[illegible]WMTNN[illegible]IV[illegible]EKIVEM
                     YSP[illegible]RQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTIL
                     KAL[illegible]PAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKC
                     FN[illegible]
                     EQTRANSPTRRELQVWGRDNNSLSEAGADRQGTVSFSFPQITLWQRPLVTIKIGGQLK
                     EALLDTGAD[illegible]TVLEEMNLPGRWKPKMI[illegible]
                     PIPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEI
                     CTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIP
                     HPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGW
                     KGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLR
                     WGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQ
                     IYAGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIA
                     EIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIATESIVIWGK
                     TPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIIGAETFYVD
                     GAANRETKLGKAGYVTNRGRQKVVSLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQY
                     ALGIIQAQPDQSESELVSQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKV
                     LFLDGIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCS
                     PGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVTTIHT
                     DNGSNFTSATVKAACWWA[illegible]VVL[illegible]MNKELKK[illegible]
                     TAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSR
                     DPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED"
BASE COUNT     1051 a   1358 c   1435 g    509 t
ORIGIN

  Query Match             84.0%;  Score 50.4;  DB 12;  Length 4353;
  Best Local Similarity   90.0%;  Pred. No. 0.00015;
  Matches   54;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggcccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              |||||| |||||||||||||||||||||||| |||||||||||| | |||| ||||||
Db        886 GACATCAGGCAGGGGCCCAAGGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 945


RESULT 12
AX306431
LOCUS       AX306431                1539 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 5 from Patent WO0188141.
ACCESSION   AX306431
VERSION     AX306431.1  GI:17645655
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (sites)
  AUTHORS   Wagner,R., Graf,M., Deml,L. and Bieler,K.
  TITLE     Synthetic gagpol genes and their uses
  JOURNAL   Patent: WO 0188141-A 5 22-NOV-2001;
            Geneart GmbH (DE)
FEATURES             Location/Qualifiers
     source          1..1539
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="sequence with optimized codons"
BASE COUNT      366 a    508 c    502 g    163 t
ORIGIN


  Query Match             78.7%;  Score 47.2;  DB 6;  Length 1539;
  Best Local Similarity   86.7%;  Pred. No. 0.0013;
  Matches   52;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggcccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              || |||| ||||||||||| |||||||| | |||||||||||  | |||| ||||||
Db        850 GATATCAGGCAGGGCCCCAAAGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 909


RESULT 13
AX306430
LOCUS       AX306430                3981 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 4 from Patent WO0188141.
ACCESSION   AX306430
VERSION     AX306430.1  GI:17645654
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (sites)
  AUTHORS   Wagner,R., Graf,M., Deml,L. and Bieler,K.
  TITLE     Synthetic gagpol genes and their uses
  JOURNAL   Patent: WO 0188141-A 4 22-NOV-2001;
            Geneart GmbH (DE)
FEATURES             Location/Qualifiers
     source          1..3981
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="sequence with optimized codons"
BASE COUNT     1010 a   1189 c   1244 g    538 t
ORIGIN


  Query Match             78.7%;  Score 47.2;  DB 6;  Length 3981;
  Best Local Similarity   86.7%;  Pred. No. 0.0012;
  Matches   52;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggcccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              || |||| |||||||||||| |||||||| | |||||||||||  | |||| ||||||
Db        850 GATATCAGGCAGGGCCCCAAAGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 909
```

```
RESULT 14
AX306429
LOCUS       AX306429                4341 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 3 from Patent WO0188141.
ACCESSION   AX306429
VERSION     AX306429.1  GI:17645653
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (sites)
  AUTHORS   Wagner,R., Grat,M., Deml,L. and Bieler,K.
  TITLE     Synthetic gaqpol genes and their uses
  JOURNAL   Patent: WO 0188141-A 3 22-NOV-2001;
            Geneart GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .4341
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="sequence with optimized codons"
BASE COUNT     1064 a    1341 c    1388 g     548 t
ORIGIN


  Query Match             78.7%;  Score 47.2;  DB 6;  Length 4341;
  Best Local Similarity   86.7%;  Pred. No. 0.0012;
  Matches   52;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 qacatcaaqcaqgqccccaaqgaqccctccqcqactacgtqqaccqctlcttcaagacc 60
           || |||| |||||||||||| ||||||||| | |||||||||||| | |||| |||||||
Db     850 GATATCAGGCAGGGCCCCAAAGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 909


RESULT 15
AX306428
LOCUS       AX306428                4343 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 2 from Patent WO0188141.
ACCESSION   AX306428
VERSION     AX306428.1  GI:17645652
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (sites)
  AUTHORS   Wagner,R., Graf,M., Deml,L. and Bieler,K.
  TITLE     Synthetic gaqpol genes and their uses
  JOURNAL   Patent: WO 0188141-A 2 22-NOV-2001;
            Geneart GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .4343
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="sequence with optimized codons"
BASE COUNT     1064 a    1341 c    1388 g     550 t
ORIGIN


  Query Match             78.7%;  Score 47.2;  DB 6;  Length 4343;
  Best Local Similarity   86.7%;  Pred. No. 0.0012;
  Matches   52;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 qacatcaaqcaqgqccccaaqgaqccctccqcqactacgtqqaccqctlcttcaagacc 60
           || |||| |||||||||||| ||||||||| | |||||||||||| | |||| |||||||
Db     850 GATATCAGGCAGGGCCCCAAAGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 909


Search completed: July  2, 2002, 21:57:33
Job time: 8950 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 23:48:09 ; Search time 291.1 Seconds
(without alignments)
353.881 Million cell updates/sec

Title: US-09-475-704-1
Perfect score: 60
Sequence: 1 gacatcaagcagccatgcaaa.....tgagccgctttttgcagagcc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_GenSeq_032802:*
1: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1980.DAT:*
2: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1981.DAT:*
3: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1982.DAT:*
4: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1983.DAT:*
5: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1984.DAT:*
6: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1985.DAT:*
7: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1986.DAT:*
8: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1987.DAT:*
9: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1988.DAT:*
10: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1989.DAT:*
11: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1990.DAT:*
12: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1991.DAT:*
13: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1992.DAT:*
14: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1993.DAT:*
15: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1994.DAT:*
16: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1995.DAT:*
17: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1996.DAT:*
18: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1997.DAT:*
19: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1998.DAT:*
20: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA1999.DAT:*
21: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA2000.DAT:*
22: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA2001.DAT:*
23: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA2002.DAT:*
24: /cgn2_6/prodata/genesseq/genesseqp_oct1/NA2003.DAT:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	21	AAA51607	HIV Gag major homo
2	60	100.0	1479	21	AAA51609	HIV synthetic Gag
3	60	100.0	1479	21	AAA51625	HIV codon optimize
4	55.2	92.0	60	21	AAA51608	HIV Gag major homo
5	55.2	92.0	1509	21	AAA51610	HIV synthetic Gag
6	55.2	92.0	1509	21	AAA51626	HIV codon optimize
7	55.2	92.0	1608	22	AAB09487	Human immunodefici
8	55.2	92.0	1914	22	AAB09488	Human immunodefici
9	55.2	92.0	2493	22	AAB09489	Human immunodefici
10	55.2	92.0	4350	22	AAB09490	Human immunodefici
11	53.6	89.3	60	21	AAA70426	HIV p55 Gag Major
12	53.6	89.3	1266	21	AAA70417	HIV Gag consensus
13	53.6	89.3	1503	21	AAZ52051	Codon optimised Hu
14	53.6	89.3	1515	21	AAA79412	Synthetic HIV Gag
15	53.6	89.3	1863	21	AAA70413	HIV Gag protease c
16	53.6	89.3	1865	21	AAA70475	HIV Gag protease c
17	53.6	89.3	1865	21	AAA70476	HIV Gag protease c
18	53.6	89.3	2841	21	AAA70415	Synthetic HIV Gag/
19	53.6	89.3	4319	21	AAA70414	HIV Gag polymerase
20	53.6	89.3	4472	21	AAA70472	HIV bicistronic co
21	53.6	89.3	4608	21	AAA70473	HIV bicistronic co
22	53.6	89.3	4689	21	AAA70471	HIV bicistronic co
23	53.6	89.3	4766	21	AAA70470	HIV bicistronic co
24	53.6	89.3	8908	21	AAZ52055	Packaging construc
25	52	86.7	4288	22	AAB20868	HIV 1 subtype (F/B
26	48.8	81.3	1482	22	AAE31286	Humanised HIV 1 ga
27	48.8	81.3	1532	19	AAV45173	Synthetic HIV gag
28	48.8	81.3	1532	22	AAE31283	Humanised HIV 1 ga
29	47.2	78.7	1539	24	AA167899	Nucleotide sequenc
30	47.2	78.7	1539	24	AA167900	Nucleotide sequenc
31	47.2	78.7	3981	24	AA167898	Nucleotide sequenc
32	47.2	78.7	4341	24	AA167897	Nucleotide sequenc
33	47.2	78.7	4343	24	AA167896	Nucleotide sequenc
34	43	71.7	4307	20	AAZ08740	HIV gag-pol SYNgp c
35	43	71.7	4307	21	AAA93972	Human immunodefici
36	43	71.7	4307	22	AAB43696	Codon optimised HI
37	43	71.7	4307	22	AAC86870	Nucleotide sequenc
38	43	71.7	4327	21	AAA93984	HIV partial leader
39	43	71.7	4353	21	AAA93983	HIV partial leader
40	43	71.7	4542	21	AAA93982	HIV complete leade
41	43	71.7	4542	22	AAB43699	Codon optimised HI
42	42.6	71.0	8969	22	AAL52320	HIV 1 non subtype
43	42.6	71.0	60	21	AAA51627	Major homology reg
44	42.6	71.0	1479	21	AAA51611	Nucleotide sequenc
45	42.6	71.0	1479	21	AAA51624	HIV 1 subtype (F/B

ALIGNMENTS

RESULT 1
AAA51607
ID AAA51607 standard; DNA; 60 BP.
XX
AC AAA51607;
XX
DT 31-OCT-2000 (first entry)
XX
DE HIV-1 Gag major homologous region nucleotides 644-703.
XX
KW expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA vaccine; HIV-1; polypeptide; cell-mediated immune response; ss.
XX
OS Human immunodeficiency virus.
XX
PN WO200039304-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31275.
XX
PR 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barnett S, Zur Megede J;
XX
DR WPI; 2000-472401/39.
XX
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal

PT especially human against HIV
XX
PS Claim 1; Page 92; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag and/or Env containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX
SQ Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 60;
Best Local Similarity 100.0%; Pred. No. 4e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcaagcagggcccccaaggaagccctccgcgactacgtggaccgcttcttccaagacc 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 gacatcaagcagggcccccaaggaagccctccgcgactacgtggaccgcttcttccaagacc 60

RESULT 2
AAA51609
ID AAA51609 standard; DNA; 1479 BP.
XX
AC AAA51609;
XX
DT 31-OCT-2000 (first entry)
XX
DE HIV synthetic Gag polynucleotide.
XX
KW Gag, expression cassette, antigenic, type C, HIV; Env; synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS Human immunodeficiency virus type C strain AF110965.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1479
FT /*tag= a
FT /product= Synthetic_Gag
FT /note= "Codon usage pattern was modified and inhibitory
FT elements (INS) and RRE sites were inactivated
FT resulting in improved expression"
XX
PN WO200039304-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31273.
XX
PR 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barnett S, Zur Megede J,
XX
DR WPI; 2000-452401/39.
DR P-PSDB; AAY96943.
XX
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX
PS Claim 2; Page 92-93; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic

CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag and/or Env containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX
SQ Sequence 1479 BP; 325 A; 526 C; 463 G; 162 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 1479;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcaagcagggcccccaaggaagccctccgcgactacgtggaccgcttcttccaagacc 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 811 gacatcaagcagggcccccaaggaagccctccgcgactacgtggaccgcttcttccaagacc 870

RESULT 3
AAA51625
ID AAA51625 standard; DNA; 1479 BP.
XX
AC AAA51625;
XX
DT 31-OCT-2000 (first entry)
XX
DE HIV codon-optimized synthetic Gag polynucleotide.
XX
KW Gag, expression cassette, antigenic, type C, HIV; Env; synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS Human immunodeficiency virus type C strain AF110965.
OS Synthetic.
XX
PN WO200039304-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31273.
XX
PR 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barnett S, Zur Megede J;
XX
DR WPI; 2000-452401/39.
XX
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX
PS Disclosure; Page 103-104; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self assemble into non infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX
SQ Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 1479;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;

```
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 gacatcaagcaggggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 903
```

RESULT 4
AAA51608
ID   AAA51608 standard; DNA; 60 BP.
XX
AC   AAA51608;
XX
DT   31-OCT-2000 (first entry)
XX
DE   HIV Gag major homology region nucleotides 841-900.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus.
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31273.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-SEP-1999;   99US-0152195.
XX
PA   (CHIR) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT   Env polypeptide and the polypeptide useful for immunizing a mammal
PT   especially human against HIV
XX
PS   Claim 1; Page 92; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic
CC   type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC   useful in DNA immunization, generation of packaging cell lines and
CC   production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC   expression cassettes exhibit increased potency for induction of
CC   cytotoxic T lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC   self-assemble into non-infectious virus-like particles which are used as
CC   a matrix for the proper presentation of an antigen entrapped or
CC   associated to the immune system of the host.
XX
SQ   Sequence 60 BP; 12 A; 24 C; 15 G; 9 T; 0 other;

```
  Query Match             92.0%;  Score 55.2;  DB 21;  Length 60;
  Best Local Similarity   95.0%;  Pred. No. 1.3e-08;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 gacatccgccaggggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
```

RESULT 5
AAA51610
ID   AAA51610 standard; DNA; 1509 BP.
XX
AC   AAA51610;
XX
DT   31-OCT-2000 (first entry)
XX
DE   HIV synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus type C strain AF110967.
OS   Synthetic.
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31273.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-SEP-1999;   99US-0152195.
XX
PA   (CHIR) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT   Env polypeptide and the polypeptide useful for immunizing a mammal
PT   especially human against HIV
XX
PS   Claim 2; Page 93; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic
CC   type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC   useful in DNA immunization, generation of packaging cell lines and
CC   production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC   expression cassettes exhibit increased potency for induction of
CC   cytotoxic T lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC   self-assemble into non-infectious virus-like particles which are used as
CC   a matrix for the proper presentation of an antigen entrapped or
CC   associated to the immune system of the host.
XX
SQ   Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 other;

```
  Query Match             92.0%;  Score 55.2;  DB 21;  Length 1509;
  Best Local Similarity   95.0%;  Pred. No. 1.9e-08;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 gacatccgccaggggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 900
```

RESULT 6
AAA51626
ID   AAA51626 standard; DNA; 1509 BP.
XX
AC   AAA51626;
XX
DT   31-OCT-2000 (first entry)
XX
DE   HIV codon optimized synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus type C strain AF110967.
OS   Synthetic.
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31273.
XX

```
PR   31-DEC-1998;   98US-0114495.
PR   01-SEP-1999;   99US-0152195
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PI   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT   Env polypeptide and the polypeptide useful for immunizing a mammal
PT   especially human against HIV
XX
PS   Disclosure; Page 104; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic
CC   type C human immunodeficiency virus (HIV) Gag or env polypeptides are
CC   useful in DNA immunization, generation of packaging cell lines and
CC   production of Gag and/or Env containing proteins. Synthetic Env and Gag
CC   expression cassettes exhibit increased potency for induction of
CC   cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV 1
CC   self-assemble into non infectious virus like particles which are used as
CC   a matrix for the proper presentation of an antigen entrapped or
CC   associated to the immune system of the host.
XX
SQ   Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 other;


  Query Match             92.0%;  Score 55.2;  DB 21;  Length 1509;
  Best Local Similarity   95.0%;  Pred. No. 1.9e-08;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqcccaaggaqccttcgcgactacgtqqaccqcttcttcaaqacc 60
           ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gacatccgccaggqcccaaggaqccttccgcgactacgtqqaccqcttcttcaaqacc 900


RESULT   7
AAD09487
ID   AAD09487 standard; DNA; 1608 BP.
XX
AC   AAD09487;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus A (HIV A) gene.
XX
KW   Human immunodeficiency virus A; HIV A; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector;
KW   MVA; cytotoxic T-lymphocyte; CTL; epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             19..1602
FT                   /*tag= a
FT                   /product= "HIV A immunogen"
XX
PN   WO200147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB04984.
XX
PR   23-DEC-1999;   99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PA   (MEDI ) MEDICAL RES COUNCIL.
PA   (IAVI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T,  Mcmichael AJ;
XX
DR   WPI; 2001-419221/44.
DR   P-PSDB; AAP04825.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T lymphocyte epitopes of HIV
PT   protein -
XX
PS   Claim 29; Fig 2A; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N terminal myristoylation, and a synthetic polypeptide comprising
CC   human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject.
CC   The invention also relates to method of stimulating anti-HIV immune
CC   response in a human subject which comprises administering one or more
CC   times an amount of nucleic acid molecule sufficient to prime an immune
CC   response to the immunogen, or else may be packaged within a delivery
CC   means, such as a modified vaccinia virus Ankara (MVA) to boost the immune
CC   response to common portion of the immunogens. The present DNA sequence
CC   encodes human immunodeficiency virus A immunogen (HIV A) fusion protein
CC   construct. HIV A immunogen consists of about 73% of gag protein fused to
CC   a string of 25 partially overlapping human CTL epitopes. The gag domain
CC   of HIV A contains p24 and p17 in an order reversed to the viral gag
CC   p17 p24-p15 polyprotein.
XX
SQ   Sequence 1608 BP; 314 A; 598 C; 458 G; 238 T; 0 other;


  Query Match             92.0%;  Score 55.2;  DB 22;  Length 1608;
  Best Local Similarity   95.0%;  Pred. No. 1.9e-08;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqcccaaggaqccttcgcgactacgtqqaccqcttcttcaaqacc 60
           ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     472 gacatccgccaggqcccaaggaqccttccgcgactacgtqqaccqcttcttcaaqacc 531


RESULT   8
AAD09488
ID   AAD09488 standard; DNA; 1914 BP.
XX
AC   AAD09488;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus IA (HIV IA) gene.
XX
KW   Human immunodeficiency virus IA; HIV IA; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector;
KW   MVA; cytotoxic T lymphocyte, CTL, epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             14..1914
FT                   /*tag= a
FT                   /product= "HIV IA immunogen"
XX
PN   WO200147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB04984.
XX
PR   23-DEC-1999;   99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
```

```
PA   (ITAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T,  Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAE94826.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T lymphocyte epitopes of HIV
PT   protein
XX
PS   Example 2; Fig 6B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject.
CC   The invention also relates to method of stimulating anti HIV immune
CC   response in a human subject which comprises administering one or more
CC   times an amount of nucleic acid molecule sufficient to prime an immune
CC   response to the immunogen, or else may be packaged within a delivery
CC   means, such as a modified vaccinia virus Ankara (MVA) to boost the immune
CC   response to common portion of the immunogens. The present DNA sequence
CC   encodes human immunodeficiency virus TA immunogen (HIV TA) fusion
CC   protein construct. HIV TA construct is based on HIV A immunogen.
CC   The HIV A immunogen consists of about 73% of gag protein fused to
CC   a string of 25 partially overlapping human CTL epitopes. The gag domain
CC   of HIV A contains p24 and p17 in an order reversed to the viral gag
CC   p17-p24-p15 polyprotein. HIV TA shares the same design rationale with
CC   HIV A construct, but additionally includes the HIV-1 clade A tat sequence
CC   positioned between the gag protein and synthetic human CTL epitopes.
XX
SQ   Sequence 1914 BP; 383 A; 728 C; 534 G; 269 T; 0 other;

  Query Match             92.0%;  Score 55.2;  DB 22;  Length 1914;
  Best Local Similarity   95.0%;  Pred. No. 1.9e-08;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaaggcc 60
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 gacatccggccaggccccaaggagcccttccgcgactacgtggaccgcttcttcaaggcc 525


RESULT 9
AAD09489
ID   AAD09489 standard; DNA; 2493 BP.
XX
AC   AAD09489;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus AeT (HIV AeT) gene.
XX
KW   Human immunodeficiency virus AeT; HIV AeT; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector;
KW   MVA; cytotoxic T-lymphocyte; CTL; epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             2179..2493
FT                   /*tag=  a
FT                   /product= "HIV AeT immunogen"
XX
PN   WO200147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB04984.
XX
PR   23-DEC-1999;   99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (ITAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T,  Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAE94827.
XX
PT   Novel immunogen for stimulating anti HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T lymphocyte epitopes of HIV
PT   protein
XX
PS   Example 2; Fig 7B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject.
CC   The invention also relates to method of stimulating anti HIV immune
CC   response in a human subject which comprises administering one or more
CC   times an amount of nucleic acid molecule sufficient to prime an immune
CC   response to the immunogen, or else may be packaged within a delivery
CC   means, such as a modified vaccinia virus Ankara (MVA) to boost the immune
CC   response to common portion of the immunogens. The present DNA sequence
CC   encodes human immunodeficiency virus AeT immunogen (HIV AeT) fusion
CC   protein construct. HIV AeT construct is based on HIV A immunogen.
CC   The HIV A immunogen consists of about 73% of gag protein fused to
CC   a string of 25 partially overlapping human CTL epitopes. The gag domain
CC   of HIV A contains p24 and p17 in an order reversed to the viral gag
CC   p17-p24-p15 polyprotein. HIV AeT shares the same design rationale with
CC   HIV A construct, but additionally includes the HIV-1 clade A tat sequence
CC   which is present on the same construct, but expressed as a separate
CC   polypeptide, by virtue of the inclusion of an internal ribosome entry
CC   site.
XX
SQ   Sequence 2493 BP; 516 A; 887 C; 685 G; 405 T; 0 other;

  Query Match             92.0%;  Score 55.2;  DB 22;  Length 2493;
  Best Local Similarity   95.0%;  Pred. No. 2e-08;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaaggcc 60
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 gacatccggccaggccccaaggagcccttccgcgactacgtggaccgcttcttcaaggcc 525


RESULT 10
AAD09490
ID   AAD09490 standard; DNA; 4350 BP.
XX
AC   AAD09490;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus PFA immunogen.
XX
KW   Human immunodeficiency virus; PFA; immunogen; anti HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector;
KW   MVA; cytotoxic T lymphocyte; CTL; epitope; ds.
```

```
XX
OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             13..4350
FT                   /*tag= a
FT                   /product= "PPA immunogen"
XX
PN   WO200147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB04984.
XX
PR   23-DEC-1999;   99GB-0030294.
PR   14-OCT-2000; 2000GB-0025224.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (ITAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA ) UNIV NAIROBI.
XX
PI   Hanke T, Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAE04828.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T lymphocyte epitopes of HIV
PT   protein -
XX
PS   Claim 40; Fig 8B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject.
CC   The invention also relates to method of stimulating anti-HIV immune
CC   response in a human subject which comprises administering one or more
CC   times an amount of nucleic acid molecule sufficient to prime an immune
CC   response to the immunogen, or else may be packaged within a delivery
CC   means, such as a modified vaccinia virus Ankara (MVA) to boost the immune
CC   response to common portion of the immunogens. The present DNA sequence
CC   encodes human immunodeficiency virus PPA immunogen fusion protein
CC   construct. PPA construct is based on HIV A immunogen. The HIV A immunogen
CC   consists of about 73% of gag protein fused to a string of 25 partially
CC   overlapping human CTL epitopes. The gag domain of HIV A contains p24 and
CC   p17 in an order reversed to the viral gag p17 p24-p15 polyprotein. PPA
CC   shares the same design rationale with HIV A construct, but additionally
CC   includes the HIV-1 clade A tat sequence, C pol protein, N pol protein,
CC   rev and nef proteins.
XX
SQ   Sequence 4350 BP; 888 A; 1630 C; 1259 G; 573 T; 0 other;


  Query Match             92.0%;  Score 55.2;  DB 22;  Length 4350;
  Best Local Similarity   95.0%;  Pred. No. 2.1e-06;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcagqqccccaaggagccctcccgcgactacgtgqaccgcttcttcaagacc 60
           ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 gacatcccgccagqgccccaaggagccctcccgcgactacgtgqaccgcttcttcaagacc 525


RESULT  11
AAA70426
ID   AAA70426 standard; DNA; 60 BP.
XX
AC   AAA70426;
XX
DT   28-NOV-2000  (first entry)
XX
DE   HIV p55 Gag Major Homology Region coding sequence.
XX
KW   HIV-1; AIDS; Gag; vaccine; expression cassette; gs.
XX
OS   Human immunodeficiency virus type 1.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..60
FT                   /*tag= a
FT                   /partial
FT                   /product= "HIV p55 Gag Major Homology Region protein"
FT                   /note= "No stop codon given"
XX
PN   WO200029302-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31245.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-DEC-1999;   99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Xu R, Mcgrath J, Srivastava I, Tian Y, Barton K, Liu H;
PI   Groer C, Selby M, Walker C;
XX
DR   WPI; 2000-452400/39.
DR   P-PSDB; AAB14214.
XX
PT   Expression cassettes encoding human immunodeficiency virus (HIV)
PT   Gag-containing polypeptide useful for vaccinating against HIV
PT   infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS   Claim 1; Page 345; 391pp; English.
XX
CC   The present invention relates to synthetic HIV Gag expression cassettes.
CC   The Gag protein of HIV is needed for the assembly of virus-like
CC   particles. In addition, the Gag protein is involved in many stages of the
CC   HIV life cycle, including assembly, virion maturation after particle
CC   release and early post entry steps in viral replication. The expression
CC   cassette may be used for the recombinant expression of HIV
CC   Gag polypeptides which may then be used to vaccinate against HIV
CC   infection and acquired immunodeficiency syndrome (AIDS). The present
CC   sequence is the coding sequence of HIV p55 Gag Major Homology Region.
CC   This sequence is located within the p24-CA sequence of Gag. Mutations in
CC   this region can severely impair particle production.
XX
SQ   Sequence 60 BP; 12 A; 24 C; 15 G; 8 T; 0 other;


  Query Match             89.4%;  Score 53.6;  DB 21;  Length 60;
  Best Local Similarity   93.8%;  Pred. No. 4.3e-08;
  Matches   56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcagqqccccaaggagccctcccgcgactacgtgqaccgcttcttcaagacc 60
           ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gacatccgccagqgccccaaggagccctcccgcgactacgtgqaccgcttctacaagacc 60


RESULT  12
AAA70417
ID   AAA70417 standard; DNA; 1268 BP.
XX
AC   AAA70417;
XX
DT   28-NOV-2000  (first entry)
XX
DE   HIV Gag common region coding sequence.
XX
```

```
KW   HIV 1; AIDS; Gag; vaccine; expression cassette; ss.
XX
OS   Human immunodeficiency virus type 1.
XX
PN   WO200039302-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31245.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-DEC-1999;   99US-0168471.
XX
PA   (CHIR-) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J, Srivastava I, Lian Y, Barton K, Liu B;
PI   Greer C, Selby M, Walker C;
XX
DR   WPI; 2000-452400/39.
XX
PT   Expression cassettes encoding the human immunodeficiency virus (HIV)
PT   Gag containing polypeptide useful for vaccinating against HIV
PT   infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS   Claim 2; Page 344; 349pp; English.
XX
CC   The present invention relates to synthetic HIV Gag expression cassettes.
CC   The Gag protein of HIV is needed for the assembly of virus like
CC   particles. In addition, the Gag protein is involved in many stages of the
CC   HIV life cycle, including assembly, virion maturation after particle
CC   release and early post entry steps in viral replication. The expression
CC   cassette may be used for the recombinant expression of HIV
CC   Gag polypeptides which may then be used to vaccinate against HIV
CC   infection and acquired immunodeficiency syndrome (AIDS). The present
CC   sequence is a common region found in Gag coding sequences: Gag protease
CC   and Gag polymerase.
XX
SQ   Sequence 1268 BP; 274 A; 449 C; 411 G; 135 T; 0 other;


  Query Match             89.3%;  Score 54.6;  DB 21;  Length 1268;
  Best Local Similarity   94.8%;  Pred. No. 5.9e-08;
  Matches   56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcagccatgcaaatgttaaaagagaccatcaacgaggaggctgcagaagtg 60
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     862 gacatccgcagccatgcaaatgttaaaagagaccatcaacgaggaggctgcagaagtg 921


RESULT  13
AAZ52051
ID   AAZ52051 standard; DNA; 1503 BP.
XX
AC   AAZ52051;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Codon optimised Human immunodeficiency virus gag coding region.
XX
KW   HIV; gag; packaging cell line; lentivirus; retroviral vector particle;
KW   pol; gene therapy; gene replacement; vaccine; biochemical reagent;
KW   codon optimisation; ds.
XX
OS   Human immunodeficiency virus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1503
FT                   /*tag= a
FT                   /product= "gag protein"
XX
PN   WO200015815-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20675.
XX
PR   11-SEP-1998;   98US-0100022.
PR   12-SEP-1998;   98US-0100063.
XX
PA   (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI   Gray JE, Mulligan RC;
XX
DR   WPI; 2000-272455/23.
DR   P-PSDB; AAY70599.
XX
PT   New packaging cell line for producing a viral accessory protein
PT   independent HIV derived retroviral vector particles, useful in gene
PT   therapy or gene replacement
XX
PS   Claim 41; Fig 8; 62pp; English.
XX
CC   The patent discloses new packaging cell line for producing a viral
CC   accessory protein independent lentivirus, preferably
CC   human immunodeficiency virus (HIV), derived retroviral vector particles.
CC   The packaging cell line comprises a mammalian cell, a retroviral RNA
CC   comprising a coding sequence for a lentivirus, preferably HIV, gagpol,
CC   where the coding sequence has been mutagenised to improve expression of
CC   the viral gagpol proteins, a second retroviral nucleotide sequence
CC   comprising the coding sequence for a heterologous envelope protein and a
CC   third retroviral nucleotide sequence comprising a DNA sequence of
CC   interest and lentivirus, preferably HIV, cis acting sequences required
CC   for packaging, reverse transcription and integration.
CC   The packaging cell lines and viral particles can be used for gene
CC   therapy or gene replacement with improved safety. They can also be used
CC   in the development and production of vaccines and biochemical reagents.
CC   The present sequence is a gag coding region of codon optimised HIV
CC   gagpol sequence. This sequence is used in the packaging
CC   cell line. Codon optimisation results in improved expression of the
CC   gagpol protein and reduces the risk of recombination between the transfer
CC   vector and gagpol mRNA.
XX
SQ   Sequence 1503 BP; 433 A; 548 C; 433 G; 201 T; 0 other;


  Query Match             89.3%;  Score 54.6;  DB 21;  Length 1503;
  Best Local Similarity   94.8%;  Pred. No. 6e-08;
  Matches   56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcagccatgcaaatgttaaaagagaccatcaacgaggaggctgcagaagtg 60
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     860 gacatccgcagccatgcaaatgttaaaagagaccatcaacgaggaggctgcagaagtg 919


RESULT  14
AAA70412
ID   AAA70412 standard; DNA; 1515 BP.
XX
AC   AAA70412;
XX
DT   28-NOV-2000  (first entry)
XX
DE   Synthetic HIV Gag expression cassette coding sequence Gag.ModSF2.
XX
KW   HIV 1; AIDS; Gag; vaccine; expression cassette; ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
PN   WO200039302-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31245.
XX
```

```
PR   31-DEC-1998;   98US-0114495.
PR   01-DEC-1999;   99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Zur Megede J,  Srivastava I,  Lian Y,  Hartog K,  Liu H;
PI   Greer C,  Selby M,  Walker C;
XX
DR   WPI; 2000-452400/39.
XX
PI   Expression cassettes encoding the human immunodeficiency virus (HIV)
PI   Gag-containing polypeptide useful for vaccinating against HIV
PI   infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS   Claim 3; Fig 7; 391pp; English.
XX
CC   The present sequence is the coding sequence of a HIV Gag expression
CC   cassette, Gag.ModSF2. The Gag protein of HIV is needed for the assembly
CC   of virus-like particles. In addition, the Gag protein is involved in
CC   many stages of the HIV life cycle, including assembly, virion maturation
CC   after particle release and early post-entry steps in viral replication.
CC   The expression cassette may be used for the recombinant expression of
CC   HIV Gag-polypeptides which may then be used to vaccinate against HIV
CC   infection and acquired immunodeficiency syndrome (AIDS).
XX
SQ   Sequence 1515 BP; 329 A; 547 C; 480 G; 159 T; 0 other;


  Query Match             89.3%;  Score 53.6;  DB 21;  Length 1515;
  Best Local Similarity   93.3%;  Pred. No. 6e-08;
  Matches   56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqcccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||   ||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     862 gacatccgccagggcccaaggagccctccgcgactacgtggaccgcttctacaagacc 921


RESULT  15
AAA70413
ID   AAA70413 standard; DNA; 1853 BP.
XX
AC   AAA70413;
XX
DT   28-NOV-2000  (first entry)
XX
DE   HIV Gag-protease expression cassette coding sequence GagProt.ModS.
XX
KW   HIV-1; AIDS; Gag-protease, vaccine, expression cassette; ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
PN   WO200039302-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31245.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-DEC-1999;   99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Zur Megede J,  Srivastava I,  Lian Y,  Hartog K,  Liu H;
PI   Greer C,  Selby M,  Walker C;
XX
DR   WPI; 2000-452400/39.
XX
PI   Expression cassettes encoding the human immunodeficiency virus (HIV)
PI   Gag-containing polypeptide useful for vaccinating against HIV
PI   infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS   Claim 5; Fig 7; 391pp; English.
XX
CC   The present sequence is the coding sequence of a HIV Gag-protease
CC   expression cassette, GagProt.ModS. The Gag protein of HIV is needed for
CC   the assembly of virus-like particles. In addition, the Gag protein is
CC   involved in many stages of the HIV life cycle, including assembly, virion
CC   maturation after particle release and early post-entry steps in viral
CC   replication. The expression cassette may be used for the recombinant
CC   expression of HIV Gag polypeptides which may then be used to vaccinate
CC   against HIV infection and acquired immunodeficiency syndrome (AIDS).
XX
SQ   Sequence 1853 BP; 421 A; 624 C; 580 G; 228 T; 0 other;


  Query Match             89.3%;  Score 53.6;  DB 21;  Length 1853;
  Best Local Similarity   93.3%;  Pred. No. 6.2e-08;
  Matches   56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqcccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||   ||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     862 gacatccgccagggcccaaggagccctccgcgactacgtggaccgcttctacaagacc 921


Search completed: July  2, 2002, 23:09:45
Job time: 8497 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 19:19:44 ; Search time 2357.28 Seconds
(without alignments)
344.562 Million cell updates/sec

Title: US-09-475-704-1
Perfect score: 60
Sequence: 1 gacatcaaaaaaaaaccccaa.....cgaaccgctttcaaggacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: qb_est1:*
10: qb_est2:*
11: qb_htc:*
12: qb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	28.6	47.7	491	9	AU031268	AU031268 AU031268
2	28.4	47.3	563	10	BG517417	BG517417 94706c12
3	28.2	47.0	585	9	AW147097	AW147097 7q70d1f07
4	28	46.7	168	10	BG050418	BG050418 EM1_53_H0
5	28	46.7	216	9	AA143927	AA143927 /EST00762
6	28	46.7	219	10	BE356040	BE356040 DC1_121_H
c 7	28	46.7	282	10	BF992820	BF992820 1020006830
8	28	46.7	327	10	BF417794	BF417794 oV1_9_E10
c 9	28	46.7	336	9	AI586607	AI586607 486049b10
10	28	46.7	347	9	AU097510	AU097510 AU097510
c 11	28	46.7	350	10	BI305785	BI305785 NL_1_L13
12	28	46.7	355	9	AU029576	AU029576 AU029576
13	28	46.7	364	13	C99276	C99276 C99276 Rice
c 14	28	46.7	364	10	T14742	T14742 05c04107 T2
15	28	46.7	364	10	BE918440	BE918440 oV1_3_G04
16	28	46.7	366	9	AU223236	AU223236 AU223236
17	28	46.7	370	10	BE595775	BE595775 PI1_54_G0
18	28	46.7	371	10	BE917843	BE917843 oV1_2_G1
19	28	46.7	372	10	BG462293	BG462293 947c4bf11
20	28	46.7	376	10	BM422819	BM422819 P1C1_2_B
21	28	46.7	381	9	AU198003	AU198003 AU198003
22	28	46.7	386	10	BI099426	BI099426 IF1_4_A
23	28	46.7	387	10	BG047940	BG047940 oV1_25_F1
24	28	46.7	389	10	BE657016	BE657016 oV2_1m_B6
25	28	46.7	391	10	BG048475	BG048475 oV1_14_A6
26	28	46.7	397	10	BG273239	BG273239 oV2_28_F6
27	28	46.7	398	6	BE954444	BE954444 PI1_14_C
28	28	46.7	402	10	BG240684	BG240684 oV1_87_A6
c 29	28	46.7	405	10	BM826219	BM826219 MES15_2_C
30	28	46.7	407	10	BE657319	BE657319 oV2_2m_F6
c 31	28	46.7	408	10	BG265180	BG265180 947c22f02
32	28	46.7	410	10	BI776652	BI776652 Ebpio7_SQ
33	28	46.7	411	10	BG415662	BG415662 HVSMEK0000
34	28	46.7	420	10	BG416623	BG416623 947c21E09
35	28	46.7	431	10	BG411913	BG411913 oV2_55_E1
36	28	46.7	441	9	AU182770	AU182770 AU182770
37	28	46.7	442	9	AU182657	AU182657 AU182657
38	28	46.7	442	10	BI778870	BI778870 EBr007_SQ
39	28	46.7	444	10	BE918452	BE918452 oV1_4_A04
40	28	46.7	445	9	AU198074	AU198074 AU198074
41	28	46.7	448	9	AU198170	AU198170 AU198170
c 42	28	46.7	449	10	BI819043	BI819043 947cc01c6
43	28	46.7	456	10	BG550542	BG550542 947d7f03
44	28	46.7	462	10	BE146208	BE146208 WHE1869_E
c 45	28	46.7	463	10	BG841753	BG841753 MES12_A1

ALIGNMENTS

RESULT 1
AU031268
LOCUS AU031268 491 bp mRNA linear EST 29-OCT-1998
DEFINITION AU031268 Rice cDNA from immature leaf including apical meristem
Oryza sativa cDNA clone E61244 2Z, mRNA sequence.
ACCESSION AU031268
VERSION AU031268.1 GI:7671758
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; oryzeae; oryza.
REFERENCE 1 (bases 1 to 491)
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from immature leaf including apical meristem
JOURNAL Unpublished (1997)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tasaki@nias.affrc.go.jp URL:http://cdna.dna.affrc.go.jp/
PROJECT: 'RGP'.
FEATURES Location/Qualifiers
source 1..491
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/clone="E61244 2Z"
/clone_lib="Rice cDNA from immature leaf including apical meristem"
/dev_stage="immature"
/note="Organ: leaf; immature leaf including apical meristem (under long day condition)"
BASE COUNT 133 a 127 c 132 g 98 t 1 others
ORIGIN

```
  Query Match             47.7%;  Score 28.6;  DB 9;  Length 491;
  Best Local Similarity   66.7%;  Pred. No. 90;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                |||||||| ||   | ||    || |||||| ||||  ||  |||   | || |||||| |
Db        183 GTCATCAAGCAGACCGCCGCCAAGACCTTCCACGACNTCGACGACTACCTCCTCAAGAAC 242


RESULT 2
BG517417
LOCUS       BG517417                 560 bp    mRNA    linear   EST 30-MAR-2001
DEFINITION  947062C12.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
            sequence.
ACCESSION   BG517417
VERSION     BG517417.1  GI:13490653
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 560)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 947062  row: C  column: 12.
FEATURES             Location/Qualifiers
     source          1..560
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="947 - 2 week shoot from Barkan lab"
                     /tissue_type="leaf and stem, including leaf base"
                     /dev_stage="2 week old seedling (3 leaves)"
                     /lab_host="XL1-Blue"
                     /note="Organ: shoot; Vector: Lambda ZAP (pBlueScript SK-);
                     Site_1: EcoRI; Site_2: XhoI; Directionally cloned using
                     Stratagene's UniZap XR cDNA cloning kit with  the 5' end
                     at the EcoRI site. The library represents 8 x 10e5
                     independent recombinant phage. The plants were greenhouse
                     grown."
BASE COUNT       84 a    206 c    174 g     96 t
ORIGIN

  Query Match             47.3%;  Score 28.4;  DB 10;  Length 560;
  Best Local Similarity   70.4%;  Pred. No. 1e+02;
  Matches   38;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          7 aagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                | ||||| |  |||| |  |||  | ||| |  ||  ||||| | ||||||||| |
Db        502 AGGCAGGACATCAAGAACCCCGGCAGCGTCAACCAGGACCCCATCTTCAAGAGC 555


RESULT 3
AW147097
LOCUS       AW147097                 585 bp    mRNA    linear   EST 03-NOV-1999
DEFINITION  707011F07.x1 707 - Mixed adult tissues from Walbot lab (FF) Zea
            mays cDNA, mRNA sequence.
ACCESSION   AW147097
VERSION     AW147097.1  GI:6194993
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 707011  row: F  column: 07.
FEATURES             Location/Qualifiers
     source          1..585
                     /organism="Zea mays"
                     /cultivar="W23"
                     /db_xref="taxon:4577"
                     /clone_lib="707 - Mixed adult tissues from Walbot lab (FF
                     )"
                     /tissue_type="tassel, kernel, silk, husk, root, leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: tassel, kernel, silk, husk, root, leaf;
                     Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
                     differentiated maize tissues from an active  Mutator
                     plant. Tissue ratio is 4:1:1:1:1:1 (tassel, kernel, silk,
                     husk, root, leaf). Unidirectionally cloned."
BASE COUNT       79 a    221 c    202 g     83 t
ORIGIN

  Query Match             47.0%;  Score 28.2;  DB 9;  Length 585;
  Best Local Similarity   68.4%;  Pred. No. 1.2e+02;
  Matches   39;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaag 57
                 | |||  | | || || |||||  | | |||||| | | | ||  ||||| ||||
Db         39 GCCATGGACCCGGACCTCAAGGCCGCGTCCGCGCCGACCTCGAGAGCTTCCTCAAG 95


RESULT 4
BG050418
LOCUS       BG050418                 168 bp    mRNA    linear   EST 25-JAN-2001
DEFINITION  FM1_53_B03.g1_A003 Floral Induced Meristem 1 (FM1) Sorghum
            propinquum cDNA, mRNA sequence.
ACCESSION   BG050418
VERSION     BG050418.1  GI:12503099
KEYWORDS    EST.
SOURCE      Sorghum propinquum.
  ORGANISM  Sorghum propinquum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 168)
  AUTHORS   Cordonnier-Pratt,M.-M., Cingle,A., Sudman,M., Marsala,C. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: floral-induced meristems
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
```

```
            High quality sequence start: 2
            High quality sequence stop: 141
            POLYA Yes.
FEATURES             Location/Qualifiers
     source          1..168
                     /organism "Sorghum propinquum"
                     /db_xref "taxon:132711"
                     /clone_lib "Floral-Induced Meristem 1 (FM1)"
                     /note "Organ: Floral-induced meristems; Vector:
                     pBluescript II from Lambda Zap II; Site_1: XhoI; Site 2:
                     EcoRI; mature plants were placed in a growth chamber for
                     15 days with 16 hr darkness  and 8 hr light (flowering is
                     induced by short-day conditions); 16 days after being
                     returned  to the greenhouse under natural long days during
                     late April/early May, meristems were  harvested The
                     library was made from poly A RNA in the cloning vector
                     lambda ZAP II. Clones to be sequenced were prepared by
                     mass excision."
BASE COUNT       39 a     58 c     50 g     21 t
ORIGIN

  Query Match             46.7%;  Score 28;  DB 10;  Length 168;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggaccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                | |||||| ||   | ||    || | |||| ||||  ||  | |   | || ||||| |
Db         86 GTCATCAAGGAGACGGCACCAAGACGTTCGACGACCTGACGACTACCTCCTCAAGAAC 145

RESULT      5
AA143927
LOCUS       AA143927                 216 bp    mRNA    linear   EST 05-DEC-1996
DEFINITION  ZES106762 Maize Leaf, Stratagene #937005 Zea mays cDNA clone
            csuh00762 5' end, mRNA sequence.
ACCESSION   AA143927
VERSION     AA143927.1  GI:1713332
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 216)
  AUTHORS   Baysdorfer,C.
  TITLE     The Maize cDNA Program
  JOURNAL   Unpublished (1993)
COMMENT     Contact: Baysdorfer C
            California State University
            Dept Biol Sci, California State Univ, Hayward, CA 94542
            Tel: 5108853459
            Fax: 5108854747
            Email: cbaysdor@baywire.csuhayward.edu
            Seq primer: SK.
FEATURES             Location/Qualifiers
     source          1..216
                     /organism "Zea mays"
                     /strain "B73"
                     /db_xref "taxon:4577"
                     /clone "csuh00762"
                     /clone_lib "Maize Leaf, Stratagene #937005"
                     /note "Vector: Uni-ZAP; Site_1: EcoRI; Site_2: XhoI; mRNA
                     isolated from illuminated leaves and sheaths of 5 week old
                     plant.  cDNA directionally cloned into vector. "
BASE COUNT       49 a     63 c     55 g     49 t
ORIGIN

  Query Match             46.7%;  Score 28;  DB 9;  Length 216;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggaccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                | ||||||| ||   | |||    || | |||| ||||  ||  |||   | || |||||| |
Db         14 GTCATCAAGGAGACGGCACCAAGACGTTCCACGACCTGACGACTACCTCCTCAAGAAC 73

RESULT      6
BE356040
LOCUS       BE356040                 219 bp    mRNA    linear   EST 20-JUL-2000
DEFINITION  DG1_121_E05.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BE356040
VERSION     BE356040.1  GI:9297597
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 219)
  AUTHORS   Cordonnier-Pratt,M. M., Gingle,A., Marsala,C., Sadman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: dark grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm  2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 2
            High quality sequence stop: 219
            POLYA No.
FEATURES             Location/Qualifiers
     source          1..219
                     /organism "Sorghum bicolor"
                     /db_xref "taxon:4558"
                     /clone_lib "Dark Grown 1 (DG1)"
                     /note "Organ: 5 day old dark grown seedlings; Vector:
                     Lambda Zap; Site_1: XhoI; Site 2: EcoRI; The library was
                     made from poly A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT       51 a     69 c     65 g     34 t
ORIGIN

  Query Match             46.7%;  Score 28;  DB 10;  Length 219;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggaccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                | ||||||| ||   | |||    || | |||| ||||  ||  |||   | || |||||| |
Db        108 GTCATCAAGGAGACGGCACCAAGACGTTCCACGACCTGACGACTACCTCCTCAAGAAC 167

RESULT      7
BI992820/c
LOCUS       BI992820                 282 bp    mRNA    linear   EST 24-OCT-2001
DEFINITION  IDP2607A02.x2 IDP  Unigene I. from Maize Genome Project Zea mays
            cDNA, mRNA sequence.
ACCESSION   BI992820
VERSION     BI992820.1  GI:16279210
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 282)
```

```
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 1020068  row: G  column: 03.
FEATURES             Location/Qualifiers
     source          1..282
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
                     /clone_lib="1020 - Unigene II from Maize Genome Project"
                     /note="This library represents the unique genes found in
                     the second round of EST sequencing at Stanford University
                     for the maize genome project. Sequences are present from
                     libraries 947 and 949.  Contigs were assembled using
                     ZmDBAssembler and 2 representatives from each contig were
                     selected for the Unigene set.  All singlets were also
                     selected."
BASE COUNT       40 a     75 c    104 g     63 t
ORIGIN

  Query Match              46.7%;  Score 28;  DB 10;  Length 282;
  Best Local Similarity    71.2%;  Pred. No. 1.2e+02;
  Matches   37;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          9 gcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||| |  |||| | |||   | ||| | ||  ||||| | ||||||||| |
Db        277 GCAGGACAGCAAGAAGCCGGCCAGCGGCAAGCAGCAGCCCATCTTCAAGAAC 226

RESULT    8
BF317794
LOCUS       BF317794                 327 bp    mRNA    linear   EST 21-NOV-2000
DEFINITION  OV1_9_F10.g1_A002 Ovary 1 (OV1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BF317794
VERSION     BF317794.1  GI:11266293
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 327)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: ovaries of varying immature stages
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 9
            High quality sequence stop: 266
            POLYA-No.
FEATURES             Location/Qualifiers
     source          1..327
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib "Ovary 1 (OV1)"
                     /note "Organ: Mix of ovaries of varying immature stages
                     from 8-week-old plants; Vector: pBluescript II from Lambda
                     Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
                     from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT       60 a     99 c     86 g     82 t
ORIGIN

  Query Match              46.7%;  Score 28;  DB 10;  Length 327;
  Best Local Similarity    66.7%;  Pred. No. 1.3e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                | ||||||| ||  | |||    || | |||| ||||  ||  |||   | || ||||| |
Db         37 GTCATCAAGGAGATCGCCACCAAGACGTTCCACGACCTCGACGACTACCTCCTCAAGAAC 96

RESULT    9
AI586607/c
LOCUS       AI586607                 336 bp    mRNA    linear   EST 07-APR-1999
DEFINITION  486019B10.x1 486 - leaf primordia cDNA library from Hake lab Zea
            mays cDNA, mRNA sequence.
ACCESSION   AI586607
VERSION     AI586607.1  GI:4572958
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 336)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave  Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 486049  row: D  column: 10.
FEATURES             Location/Qualifiers
     source          1..336
                     /organism "Zea mays"
                     /cultivar "B73"
                     /db_xref="taxon:4577"
                     /clone_lib "486 - leaf primordia cDNA library from Hake
                     lab"
                     /tissue_type "leaf primordia"
                     /dev_stage="P7-P11 leaf"
                     /lab_host "E.coli XLI-Blue MFR'"
                     /note "organ: shoot, Vector: Lambda zap, Hake lab cDNA
                     library."
BASE COUNT       82 a     91 c    101 g     62 t
ORIGIN

  Query Match              46.7%;  Score 28;  DB 9;  Length 336;
  Best Local Similarity    66.7%;  Pred. No. 1.3e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                | ||||||| ||  | |||    || | |||| ||||  ||  |||   | || ||||| |
Db        313 GTCATCAAGGAGATCGCCACCAAGACGTTCCACGACCTCGACGACTACCTCCTCAAGAAC 254

RESULT   10
AW097510
LOCUS       AW097510                 347 bp    mRNA    linear   EST 30-JUN-2000
```

```
DEFINITION  AU097510 Rice shoot Oryza sativa cDNA clone S5505, mRNA sequence.
ACCESSION   AU097510
VERSION     AU097510.1  GI:8860192
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 347)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from etiolated shoot (2000)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT  'RGP'.
            S5505_8Z.
FEATURES             Location/Qualifiers
     source          1..347
                     /organism="Oryza sativa"
                     /strain="Nipponbare, sub_species japonica"
                     /db_xref="taxon:4530"
                     /clone="S5505"
                     /clone_lib="Rice shoot"
                     /note="Etiolated shoot (8 days old)"
BASE COUNT       86 a     95 c     95 g     70 t      1 others
ORIGIN


  Query Match             46.7%;  Score 28;  DB 9;  Length 347;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
               | |||||| ||  | ||    || ||||:| ||||  ||  |||  | || |||||| |
Db         63 GTCATCAAGGAGACGGCGGCAAGACCTTCCACGACCTCGACGACTACCTCCTCAAGAAC 122


RESULT 11
BI305785/c
LOCUS       BI305785               350 bp    mRNA    linear   EST 20-JUL-2001
DEFINITION  NL_1_L13 Drought stress (leaf) Oryza sativa cDNA clone NL_1_L13 ...,
            mRNA sequence.
ACCESSION   BI305785
VERSION     BI305785.1  GI:14981094
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   Reddy,A.R., Ramakrishna,W., Chandrasekhar,A., Sarabhashan,L.,
            Ravindrababu,P. and Bennetzen,J.L.
  TITLE     Novel EST enrichment with normalized cDNA libraries from drought
            stressed rice (Oryza sativa L.cv Nagina 22)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Reddy AR
            Department of Plant Sciences, School of Life Sciences
            University of Hyderabad
            P.O. Central University, Hyderabad-500 046, A.P, India
            Tel: 0091 40 3010265
            Fax: 0091 40 3010145
            Email: arjulsl@uohyd.ernet.in
            Insert Length: 350   Std Error: 0.00
            Plate: 1  row: L  column: 13
            Seq primer: CTAAAACGACGGCCAGTG.
FEATURES             Location/Qualifiers
     source          1..350
                     /organism="Oryza sativa"
                     /cultivar="Nagina 22 (indica sub sp)"
                     /db_xref="taxon:4530"
                     /clone="NL_1_L13"
                     /clone_lib="Drought stress (leaf)"
                     /tissue_type="Entire leaf tissue"
                     /dev_stage="35 day old seedlings"
                     /note="Organ: Leaf; Vector: 17I3Pac; ESTs from normalized
                     leaf cDNA Library from drought stressed seedlings."
BASE COUNT       73 a    101 c    103 g     71 t
ORIGIN


  Query Match             46.7%;  Score 28;  DB 10;  Length 350;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
               | |||||| ||  | ||    || ||||| ||||  ||  |||  | || |||||| |
Db        270 GTCATCAAGGAGACGGCGGCAAGACCTTCCACGACCTCGACGACTACCTCCTCAAGAAC 211


RESULT 12
AU029576
LOCUS       AU029576               355 bp    mRNA    linear   EST 19-OCT-1998
DEFINITION  AU029576 Rice panicle shorter than 3cm Oryza sativa cDNA clone
            E31104_6Z, mRNA sequence.
ACCESSION   AU029576
VERSION     AU029576.1  GI:3762824
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 355)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from panicle
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT  'RGP'.
FEATURES             Location/Qualifiers
     source          1..355
                     /organism="Oryza sativa"
                     /strain="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="E31104_6Z"
                     /clone_lib="Rice panicle shorter than 3cm"
                     /dev_stage="shorter than 3cm"
                     /note="Organ: panicle"
BASE COUNT       81 a     97 c     86 g     67 t      1 others
ORIGIN


  Query Match             46.7%;  Score 28;  DB 9;  Length 355;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
               | |||||| ||  | |:    || ||||| ||||  ||  |||  : || |||||| |
Db         69 GTCATCAAGGAGACGGCGGCAAGACCTTCCACGACCTCGACGACTACCTCCTCAAGAAC 128


RESULT 13
C99276
LOCUS       C99276                 363 bp    mRNA    linear   EST 19-OCT-1998
```

```
DEFINITION  C99276 Rice panicle at ripening stage Oryza sativa cDNA clone
            E10581_6Z, mRNA sequence.
ACCESSION   C99276
VERSION     C99276.1  GI:3762028
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 363)
  AUTHORS   Sasaki,T.
  TITLE     Rice cDNA from panicle at ripening stage
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2 1 2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@nbr affrc go jp URL-http://rgp dna affrc go jp/
            PROJECT ='RGP'.
FEATURES             Location/Qualifiers
     source          1. .363
                     /organism "Oryza sativa"
                     /strain="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="E10581_6Z"
                     /clone_lib="Rice panicle at ripening stage"
                     /dev_stage="ripening stage"
                     /note="Organ: panicle; Rice cDNA from panicle at ripening
                     stage"
BASE COUNT       83 a    108 c    100 g     68 t      4 others
ORIGIN


  Query Match           46.7%;  Score 28;  DB 10;  Length 363;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqccccaaggaqccctteeqeqactacgtqqaccgcttcttcaaqaqcc 60
           | ||||||| ||  | ||    || |||||| |||| ||  ||| | || |||||| |
Db      80 GTCATCAAGGAGACCGCCGCCAAGACCTCCACGACCTCGACGACTACCTCCTCCAAGAAC 139


RESULT  14
T14732/c
LOCUS       T14732                   364 bp    mRNA    linear   EST 17-OCT-1996
DEFINITION  05c04f07-f21 membrane-free polysomes from endosperm Zea mays cDNA
            clone 05c04f07 3' end, mRNA sequence.
ACCESSION   T14732
VERSION     T14732.1  GI:440711
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 364)
  AUTHORS   Shen,B., Carneiro,N., Torres Jerez,I., Stevenson,E., McCreery,T.,
            Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habben,J. and
            Larkins,B.
  TITLE     Partial sequencing and mapping of clones from two maize cDNA
            libraries
  JOURNAL   Plant Mol. Biol. 26, 1085-1101 (1994)
  MEDLINE   95111093
COMMENT     Contact: The Maize cDNA Project

            Helentjaris TG (primary contact)
            Dept. of Plant Sciences
            University of Arizona
            Dept. of Plant Sciences, University of Arizona  Tucson, AZ,85721
            ph: 602-6218-746
            fax: 602-621-7186
            E-mail: helnjars@ccit.arizona.edu

            Chris Baysdorfer
            Department of Biological Sciences, School of Science
            California State University, Hayward
            Hayward, CA 94542
            ph: 510-881 3459
            fax: 510 727 2035
            E-mail: cbaysdor@sl.csuhayward.edu

            Rob Ferl
            Interdisciplinary Center for Biotechnology Research
            DNA Sequencing Core
            University of Florida
            P.O. Box 100695
            Gainesville, FL 32611-0695
            ph: 904-392-1928, ext. 301
            fax: 904-392-4072
            E-mail: robferl@nervm.nerdc.ufl.edu
            Seq primer: F21.
FEATURES             Location/Qualifiers
     source          1. .364
                     /organism="Zea mays"
                     /strain "W64A2"
                     /db_xref "taxon:4577"
                     /clone="05c04f07"
                     /clone_lib "membrane-free polysomes from endosperm"
                     /lab_host="DH10B"
                     /note "Vector: Ziplox; Site_1: SalI; Site_2: NotI; ds cDNA
                     was prepared from oligo-dT selected mRNA by priming with a
                     NotI oligo dT oligomer and then adding the second strand
                     to RNase nicked DNA:RNA hybrid with DNA Pol I. SalI
                     adaptors were added to the ends, the ds-cDNAs were then
                     digested with NotI and size-selected. These were
                     directionally-cloned into the Ziplox plant vector, excised
                     as plasmids, and then individually analyzed. "
BASE COUNT       89 a     98 c    107 g     68 t      2 others
ORIGIN


  Query Match           46.7%;  Score 28;  DB 10;  Length 364;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqccccaaggaqccctteeqeqactacgtqqaccgcttcttcaaqaqcc 60
           | ||||||| ||  | ||    || |||||| |||| ||  ||| | || |||||| |
Db     325 GTCATCAAGGAGACCGCCGCCAAGACCTCCACGACCTCGACGACTACCTCCTCCAAGAAC 266


RESULT  15
BE918440
LOCUS       BE918440                 364 bp    mRNA    linear   EST 29-SEP-2000
DEFINITION  OVI_3_G04.g1_A002 Ovary 1 (OVI) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BE918440
VERSION     BE918440.1  GI:10421546
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 364)
  AUTHORS   Cordonnier-Pratt,M. M., Cinalo,A., Marsala,C., Suhman,M. and Pratt
            ,L.H.
  TITLE     An EST Database from Sorghum: ovaries of varying immature stages
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602 7271, USA
            Tel: 706 542 1860
```

```
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: Poly1Mix
            High quality sequence start: 8
            High quality sequence stop: 358
            POLYA No.
FEATURES             Location/Qualifiers
     source          1..364
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="ovary 1 (OV1)"
                     /note="Organ: Mix of ovaries of varying immature stages
                     from 8-week-old plants; Vector: pBluescript II from Lambda
                     Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
                     from poly A RNA in the cloning vector Lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT       66 a    107 c    103 g     88 t
ORIGIN


  Query Match             46.7%;  Score 28;  DB 10;  Length 364;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcaggtccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                | ||||||| ||  | |||   || | ||| ||||  ||  |||  | || |||||| |
Db         75 GTCATCAAGGAGACCGCTACCAAGACGTTCCACGACCTCGACGACTACCTCCTCAAGAAC 134


Search completed: July  2, 2002, 21:25:55
Job time: 7552 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 19:45:13 ; Search time 63.61 Seconds
(without alignments)
231.693 Million cell updates/sec

Title: US-09-475-704-1
Perfect score: 60
Sequence: 1 gacatcaagcagggccccaa..........tggaccgcttcttcaagacc 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 683533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: [illegible]
2: [illegible]
3: [illegible]
4: [illegible]
5: [illegible]
6: [illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	35	58.3	9207	3	US-08-388-353-800	Sequence 800, App
2	35	58.3	9207	3	US-08-488-551B-800	Sequence 800, App
3	33.4	55.7	718	6	5204259-2	Patent No. 5204259
4	33.4	55.7	793	6	5204259-4	Patent No. 5204259
5	33.4	55.7	796	6	5204259-6	Patent No. 5204259
6	33.4	55.7	845	1	US-08-589-446-3	Sequence 3, Appli
7	33.4	55.7	845	1	US-08-444-882-3	Sequence 3, Appli
8	33.4	55.7	845	1	US-08-389-459A-3	Sequence 3, Appli
9	33.4	55.7	845	3	US-08-987-867A-3	Sequence 3, Appli
10	33.4	55.7	871	6	5204259-10	Patent No. 5204259
11	33.4	55.7	1921	6	5204259-8	Patent No. 5204259
12	33.4	55.7	5362	3	US-08-463-210-5	Sequence 5, Appli
13	33.4	55.7	7228	2	US-08-850-049-128	Sequence 128, App
14	33.4	55.7	7228	2	US-08-850-049-129	Sequence 129, App
15	33.4	55.7	7228	2	US-08-850-478-128	Sequence 128, App
16	33.4	55.7	7228	2	US-08-850-478-129	Sequence 129, App
17	33.4	55.7	7228	4	US-09-414-117-128	Sequence 128, App
18	33.4	55.7	7228	4	US-09-414-117-129	Sequence 129, App
19	33.4	55.7	7399	2	US-08-418-848A-9	Sequence 9, Appli
20	33.4	55.7	9709	2	US-08-188-583-5	Sequence 5, Appli
21	33.4	55.7	9709	3	US-08-388-353-1	Sequence 1, Appli
22	33.4	55.7	9709	3	US-08-488-551B-1	Sequence 1, Appli
23	33.4	55.7	9739	1	US-08-022-835-1	Sequence 1, Appli
24	33.4	55.7	9739	1	US-08-388-809-1	Sequence 1, Appli
25	33.4	55.7	9739	2	US-08-647-714-1	Sequence 1, Appli
26	33.4	55.7	9746	1	US-08-022-835-3	Sequence 3, Appli
27	33.4	55.7	9746	1	US-08-388-809-3	Sequence 3, Appli
28	33.4	55.7	9746	2	US-08-647-714-3	Sequence 3, Appli
29	33.4	55.7	12494	4	US-08-935-312-13	Sequence 13, Appl
30	33.4	55.7	12494	4	US-08-848-760B-33	Sequence 33, Appl
31	33.4	55.7	15581	3	US-08-646-538-35	Sequence 35, Appl
32	33.4	55.7	15581	4	US-09-503-222-35	Sequence 35, Appl
33	31.8	53.0	1014	1	US-07-648-796A-9	Sequence 9, Appli
34	31.8	53.0	1377	1	US-07-648-796A-13	Sequence 13, Appl
35	31.8	53.0	2313	1	US-07-648-796A-16	Sequence 16, Appl
36	31.8	53.0	2400	1	US-07-648-796A-15	Sequence 15, Appl
37	31.8	53.0	3807	2	US-08-417-216A-78	Sequence 78, Appl
38	31.8	53.0	9737	2	US-08-944-449-7	Sequence 7, Appli
39	31.2	52.0	10163	2	US-08-659-251-1	Sequence 1, Appli
40	31.2	52.0	10163	4	US-09-256-490-1	Sequence 1, Appli
41	31.2	52.0	10163	5	PCT-US96-11445-1	Sequence 1, Appli
42	30.2	50.3	1314	3	US-08-392-794A-1	Sequence 1, Appli
43	30.2	50.3	1733	1	US-08-470-202-57	Sequence 57, Appl
44	30.2	50.3	1733	1	US-08-470-202-58	Sequence 58, Appl
45	30.2	50.3	1733	1	US-08-471-770-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-08-388-353-800
; Sequence 800, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 800:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-388-353-800

Query Match 58.3%; Score 35; DB 3; Length 9207;
Best Local Similarity 74.6%; Pred. No. 0.0054;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy       1 gacatcaagcaggqccccaaqgaqccctccgcgactacgtggaccgcttcttcaagac 59
           ||||| || || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db    1458 GACATAAAACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1516


RESULT   2
US-08-488-551B-800
; Sequence 800, Application US/08488551B
; Patent No. 6015661
;  GENERAL INFORMATION:
;     APPLICANT:  Nicholas J. Deacon
;     APPLICANT:  Dale A. McPhee
;     APPLICANT:  David Cooper
;     TITLE OF INVENTION:  NON PATHOGENIC STRAINS OF HIV-1
;     NUMBER OF SEQUENCES:  841
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;       STREET:  400 GARDEN CITY PLAZA
;       CITY:  GARDEN CITY
;       STATE:  NEW YORK
;       COUNTRY:  U.S.A.
;       ZIP:  11530-0299
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/488,551B
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PM3864 (AU)
;       FILING DATE:  14-FEB-1994
;       APPLICATION NUMBER:  PM4002 (AU)
;       FILING DATE:  21-FEB-1994
;       APPLICATION NUMBER:  PN0284 (AU)
;       FILING DATE:  23-DEC-1994
;       APPLICATION NUMBER:  US 08/388,353
;       FILING DATE:  14-FEB-1995
;       APPLICATION NUMBER:  PN3021/95
;       FILING DATE:  17-MAY-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  FRANK S. DIGIGLIO
;       REFERENCE/DOCKET NUMBER:  5606%
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 742-4343
;       TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  800:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9207 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-488-551B-800


  Query Match             58.3%;  Score 35;  DB 3;  Length 9207;
  Best Local Similarity   74.6%;  Pred. No. 0.0054;
  Matches   44;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqccccaaqgaqccctccgcgactacgtggaccgcttcttcaagac 59
           ||||| || || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db    1458 GACATAAAACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1516


RESULT   3
5204259-2
;Patent No. 5204259
;     APPLICANT: HELING,TORSTEN B.;CREVIN, HAKAN;NUNN,
;MICHAEL F.
;     TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
;     NUMBER OF SEQUENCES: 20
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/07/344,237
;       FILING DATE: 26-APR-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 191,299
;       FILING DATE: 06-MAY-1988
;       APPLICATION NUMBER:191,229
;       FILING DATE: 06-MAY-1988
;       APPLICATION NUMBER: 206,499
;       FILING DATE: 13-JUN-1988
;       APPLICATION NUMBER: 258,016
;       FILING DATE: 14-OCT-1988
;SEQ ID NO:2:
;       LENGTH:718
5204259-2


  Query Match             55.7%;  Score 33.4;  DB 6;  Length 718;
  Best Local Similarity   72.9%;  Pred. No. 0.014;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqccccaaqgaqccctccgcgactacgtggaccgcttcttcaagac 59
           ||||| |  || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db     472 gacataaqacaaggaccaaaggaaccctttagagactatgtagaccggttctataaaac 530


RESULT   4
5204259-4
;Patent No. 5204259
;     APPLICANT: HELING,TORSTEN B.;CREVIN, HAKAN;NUNN,
;MICHAEL F.
;     TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
;     NUMBER OF SEQUENCES: 20
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/07/344,237
;       FILING DATE: 26-APR-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 191,299
;       FILING DATE: 06-MAY-1988
;       APPLICATION NUMBER:191,229
;       FILING DATE: 06-MAY-1988
;       APPLICATION NUMBER: 206,499
;       FILING DATE: 13-JUN-1988
;       APPLICATION NUMBER: 258,016
;       FILING DATE: 14-OCT-1988
;SEQ ID NO:4:
;       LENGTH:793
5204259-4


  Query Match             55.7%;  Score 33.4;  DB 6;  Length 793;
  Best Local Similarity   72.9%;  Pred. No. 0.014;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gacatcaagcaggqccccaaqgaqccctccgcgactacgtggaccgcttcttcaagac 59
           ||||| |  || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db     472 gacataaqacaaggaccaaaggaaccctttagagactatgtagaccggttctataaaac 530


RESULT   5
5204259-6
;Patent No. 5204259
;     APPLICANT: HELING,TORSTEN B.;CREVIN, HAKAN;NUNN,
;MICHAEL F.
;     TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
;     NUMBER OF SEQUENCES: 20
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/07/344,237
;       FILING DATE: 26-APR-1989
;     PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER: 191,299
;       FILING DATE: 06 MAY 1988
;       APPLICATION NUMBER:191,229
;       FILING DATE: 06 MAY 1988
;       APPLICATION NUMBER: 206,199
;       FILING DATE: 13 JUN 1988
;       APPLICATION NUMBER: 258,016
;       FILING DATE: 14 OCT 1988
;SEQ ID NO:6:
;       LENGTH:796
5204259-6

  Query Match             55.7%;  Score 33.4;  DB 6;  Length 796;
  Best Local Similarity   72.9%;  Pred. No. 0.014;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagac 59
              ||||| |  || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db        472 gacataagacaaggaccaaaggaacccttt agagactatgtagaccggttctataaaac 530

RESULT 6
US-08-589-446-3
; Sequence 3, Application US/08589446
; Patent No. 5614413
;  GENERAL INFORMATION:
;    APPLICANT:  Morrow, Casey D.
;    TITLE OF INVENTION:  ENCAPSIDATED POLIOVIRUS NUCLEIC
;    TITLE OF INVENTION:  ACID AND METHODS OF MAKING AND
;    TITLE OF INVENTION:  USING SAME
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE STREET, SUITE 510
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/589,446
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/087,009
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Geary III, William C.
;      REGISTRATION NUMBER:  31,359
;      REFERENCE/DOCKET NUMBER:  UAG-004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  845 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20..845
US-08-589-446-3

  Query Match             55.7%;  Score 33.4;  DB 1;  Length 845;
  Best Local Similarity   72.9%;  Pred. No. 0.014;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagac 59
              ||||| |  || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db        482 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 540

RESULT 7
US-08-444-882-3
; Sequence 3, Application US/08444882
; Patent No. 5622705
;  GENERAL INFORMATION:
;    APPLICANT:  Morrow, Casey D.
;    TITLE OF INVENTION:  ENCAPSIDATED POLIOVIRUS NUCLEIC
;    TITLE OF INVENTION:  ACID AND METHODS OF MAKING AND
;    TITLE OF INVENTION:  USING SAME
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE STREET, SUITE 510
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,882
;      FILING DATE:  19-MAY-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/087,009
;      FILING DATE:  01-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Geary III, William C.
;      REGISTRATION NUMBER:  31,359
;      REFERENCE/DOCKET NUMBER:  UAG-004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  845 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20..845
US-08-444-882-3

  Query Match             55.7%;  Score 33.4;  DB 1;  Length 845;
  Best Local Similarity   72.9%;  Pred. No. 0.014;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagac 59
              ||||| |  || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db        482 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 540

RESULT 8
US-08-389-459A-3
; Sequence 3, Application US/08389459A
; Patent No. 5817512
;  GENERAL INFORMATION:
;    APPLICANT:  Morrow, Casey D. and Porter, Donna, C.
;    TITLE OF INVENTION:  ENCAPSIDATED RECOMBINANT POLIOVIRUS
;    TITLE OF INVENTION:  NUCLEIC ACID AND METHODS OF MAKING AND
```

```
;    TITLE OF INVENTION:  USING SAME
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE STREET, SUITE 510
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/389,459A
;      FILING DATE:  15-FEB-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/087,009
;      FILING DATE:  01-JUL-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Silveri, Jean M.
;      REGISTRATION NUMBER:  39,030
;      REFERENCE/DOCKET NUMBER:  UAG-004CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  845 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20..845
US-08-389-459A-3


  Query Match             55.7%;  Score 33.4;  DB 1;  Length 845;
  Best Local Similarity   72.9%;  Pred. No. 0.014;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 qacatcaagcaqqqccccaaggaqccctccqcgactacgtqqaccqcttcttcaaqac 59
           ||||| |  || || || ||||| |||||  | |||| || ||||| |||| || ||
Db     482 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 540


RESULT   9
US-08-987-867A-3
; Sequence 3, Application US/08987867A
; Patent No. 6063384
;  GENERAL INFORMATION:
;    APPLICANT:  C. Morrow et al.
;    TITLE OF INVENTION:  ENCAPSIDATED RECOMBINANT VIRAL
;    TITLE OF INVENTION:  NUCLEIC ACID AND METHODS OF MAKING AND
;    TITLE OF INVENTION:  USING SAME
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  28 STATE STREET
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,867A
;      FILING DATE:  09-DEC-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/087,009
;      FILING DATE:  01-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Myers, Louis
;      REGISTRATION NUMBER:  35,965
;      REFERENCE/DOCKET NUMBER:  UAP-004CPDV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 742-4214
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  845 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20..845
US-08-987-867A-3


  Query Match             55.7%;  Score 33.4;  DB 4;  Length 845;
  Best Local Similarity   72.9%;  Pred. No. 0.014;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 qacatcaagcaqqqccccaaggaqccctccqcgactacgtqqaccqcttcttcaaqac 59
           ||||| |  || || || ||||| |||||  | |||| || ||||| |||| || ||
Db     482 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 540


RESULT   10
5204259-10
;Patent No. 5204259
;    APPLICANT: HELTING,TORCHEN R.;CREVIN, HAKAN;NUNN,
;MICHAEL F.
;    TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
;    NUMBER OF SEQUENCES: 20
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/244,237
;      FILING DATE: 26-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 191,299
;      FILING DATE: 06-MAY-1988
;      APPLICATION NUMBER:191,229
;      FILING DATE: 06-MAY-1988
;      APPLICATION NUMBER: 206,499
;      FILING DATE: 13-JUN-1988
;      APPLICATION NUMBER: 258,016
;      FILING DATE: 14-OCT-1988
;SEQ ID NO:10:
;      LENGTH:871
5204259-10


  Query Match             55.7%;  Score 33.4;  DB 6;  Length 871;
  Best Local Similarity   72.9%;  Pred. No. 0.014;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 qacatcaagcaqqqccccaaggaqccctccqcgactacgtqqaccqcttcttcaaqac 59
           ||||| |  || || || ||||| |||||  | |||| || ||||| |||| || ||
Db     172 qacatgaagcaaggaccaaaggaaccctttagagactattgtagaccggttctataaaac 530


RESULT   11
5204259-8
;Patent No. 5204259
```

; APPLICANT: HELLING,TORSTEN B.;CREVIN, HAKAN;NUNN,
;MICHAEL F.
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/344,237
; FILING DATE: 26-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 191,299
; FILING DATE: 06-MAY-1988
; APPLICATION NUMBER:191,229
; FILING DATE: 06-MAY-1988
; APPLICATION NUMBER: 206,499
; FILING DATE: 13-JUN-1988
; APPLICATION NUMBER: 258,916
; FILING DATE: 14-OCT-1988
;SEQ ID NO:8:
; LENGTH:1021
5204259-8

```
  Query Match             55.7%;  Score 33.4;  DB 6;  Length 1021;
  Best Local Similarity   72.9%;  Pred. No. 0.015;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccggttcttcaaagac 59
              |||||  |    | ||  |  | |||  |||||   |  ||||| || ||||| ||||  || ||
Db        472 gacataagacaaggaccaaaggaacccttagagactatgtagaccggttctataaaac 530
```

RESULT 12
US-08-463-210-5
; Sequence 5, Application US/08463210
; Patent No. 6001977
; GENERAL INFORMATION:
; APPLICANT: CHANG, Nancy T.
; APPLICANT: GALLO, Robert C.
; APPLICANT: WONG-STAAL, Flossie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,210
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/693,866
; FILING DATE: 23-JAN-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/659,339
; FILING DATE: 10-OCT-1984
; ATTORNEY/AGENT INFORMATION:
; NAME: Seranian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4193US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5362 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: HTLV III
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..5362
; OTHER INFORMATION: /standard_name= "Clone BH5"
; OTHER INFORMATION: /note= "Corresponds to nucleotide positions 222 to
; OTHER INFORMATION: 5585 in Figure 3 of US 06/693,866 (parent)"
US-08-463-210-5

```
  Query Match             55.7%;  Score 33.4;  DB 3;  Length 5362;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatcaagcagggccccaaggagcccttccgcgactacgtggaccggttcttcaaagac 59
              |||||  |    | ||  || ||||| |||||   |  ||||| || |  ||| ||||  || ||
Db        960 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1018
```

RESULT 13
US-08-850-049-128
; Sequence 128, Application US/08850049
; Patent No. 5965726
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT:
; APPLICANT:
; TITLE OF INVENTION: METHOD OF ELIMINATING
; TITLE OF INVENTION: INHIBITORY/INSTABILITY REGIONS OF mRNA
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,049
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/330,478
; FILING DATE: 26-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/02908
; FILING DATE: 29-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/858,747
; FILING DATE: 27-MAR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MORRY, MARY J.
; REGISTRATION NUMBER: 34,398
; REFERENCE/DOCKET NUMBER: 2026-4096US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:

```
;       LENGTH:  7228 BASE PAIRS
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
US-08-850-049-128


  Query Match             55.7%;  Score 34.4;  DB 2,  Length 7228;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gacatcaaqcaqqqccccdaqgaqccctccqcqactacatqqaccqttcttcaaqac 59
           ||||| |  || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db    1578 GACATAAGACAAGGACCAAAGGAACCCTTAGAGACTATGTAGACCGGTTCTATAAAAC 1636


RESULT  14
US-08-850-049-129
; Sequence 129, Application US/08850049
; Patent No. 5965726
;  GENERAL INFORMATION:
;     APPLICANT:
;     APPLICANT:
;     APPLICANT:
;     TITLE OF INVENTION:  METHOD OF ELIMINATING
;     TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;     NUMBER OF SEQUENCES:  130
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/850,049
;       FILING DATE:  02-MAY-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/850,478
;       FILING DATE:  26 OCT 1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/02908
;       FILING DATE:  29-MAR-1993
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/858,747
;       FILING DATE:  27-MAR-1992
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  MORRY, MARY J.
;       REGISTRATION NUMBER:  34,398
;       REFERENCE/DOCKET NUMBER:  2926-4006US1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)758-4800
;       TELEFAX:  (212)751-6849
;   INFORMATION FOR SEQ ID NO:  129:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7228 BASE PAIRS
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
US-08-850-049-129


  Query Match             55.7%;  Score 34.4;  DB 2;  Length 7228;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gacatcaaqcaqqqccccdaqgaqccctccqcqactacatqqaccqttcttcaaqac 59
           ||||| |  || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db    1578 GACATAAGACAAGGACCAAAGGAACCCTTAGAGACTATGTAGACCGGTTCTATAAAAC 1636


RESULT  15
US-08-050-478-128
; Sequence 128, Application US/08050478
; Patent No. 5972596
;  GENERAL INFORMATION:
;     APPLICANT:
;     APPLICANT:
;     APPLICANT:
;     TITLE OF INVENTION:  METHOD OF ELIMINATING
;     TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;     NUMBER OF SEQUENCES:  130
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/050,478
;       FILING DATE:  26-OCT-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/02908
;       FILING DATE:  29-MAR-1993
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/858,747
;       FILING DATE:  27-MAR-1992
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  MORRY, MARY J.
;       REGISTRATION NUMBER:  34,398
;       REFERENCE/DOCKET NUMBER:  2926-4006US1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)758-4800
;       TELEFAX:  (212)751-6849
;   INFORMATION FOR SEQ ID NO:  128:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7228 BASE PAIRS
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
US-08-050-478-128


  Query Match             55.7%;  Score 34.4;  DB 2;  Length 7228;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gacatcaaqcaqqqccccdaqgaqccctccqcqactacatqqaccqttcttcaaqac 59
           ||||| |  || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db    1578 GACATAAGACAAGGACCAAAGGAACCCTTAGAGACTATGTAGACCGGTTCTATAAAAC 1636


Search completed: July  2, 2002, 21:58:57
Job time: 8024 sec
```

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 21:57:33 ; Search time 1898.13 Seconds
(without alignments)
661.489 Million cell updates/sec

Title: US-09-475-704-2
Perfect score: 60
Sequence: 1 gacatccgccaggcccaa..........tggaccgcttcttcaagacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	60	100.0	1608	6	AX188558	AX188558 Sequence
2	60	100.0	1914	6	AX188560	AX188560 Sequence
3	60	100.0	2493	6	AX188562	AX188562 Sequence
4	60	100.0	4350	6	AX188564	AX188564 Sequence
5	58.4	97.3	1509	12	AF201927	AF201927 Synthetic
6	58.4	97.3	1847	12	AF202464	AF202464 Synthetic
7	58.4	97.3	1847	12	AF202465	AF202465 Synthetic
8	48.8	81.3	1548	12	AF287354	AF287354 Synthetic
9	48.8	81.3	4352	12	AF287352	AF287352 Synthetic
10	48.8	81.3	4353	12	AF287353	AF287353 Synthetic
11	47.8	79.7	4307	6	AX019132	AX019132 Sequence
12	47.8	79.7	4307	6	AX035453	AX035453 Sequence
13	47.8	79.7	4307	6	AX056836	AX056836 Sequence
14	47.8	79.7	4307	6	AX283597	AX283597 Sequence
15	47.8	79.7	4307	6	AX283607	AX283607 Sequence
16	47.8	79.7	4327	6	AX035465	AX035465 Sequence
17	47.8	79.7	4353	6	AX035464	AX035464 Sequence
18	47.8	79.7	4642	6	AX035463	AX035463 Sequence
19	47.8	79.7	9772	6	AX283584	AX283584 Sequence
20	47.2	78.7	4288	6	AX149648	AX149648 Sequence
21	45.6	76.0	1539	6	AX306431	AX306431 Sequence
22	45.6	76.0	3981	6	AX306430	AX306430 Sequence
23	45.6	76.0	4343	6	AX306429	AX306429 Sequence
24	45.6	76.0	4343	6	AX306428	AX306428 Sequence
25	38.2	63.7	693	14	AB034323	AB034323 Human imm
26	38.2	63.7	9016	14	AF110974	AF110974 HIV 1 iso
27	38.2	63.7	9016	14	AF110975	AF110975 HIV 1 iso
28	38.2	63.7	9012	14	AF110973	AF110973 HIV 1 iso
29	38.2	63.7	9064	14	AF443094	AF443094 HIV 1 iso
30	36.6	61.0	576	14	HIV224196	AJ224196 Human imm
31	36.6	61.0	646	14	HIM274561	AJ274561 Human imm
32	36.6	61.0	693	14	AB034321	AB034321 Human imm
33	36.6	61.0	693	14	AB034322	AB034322 Human imm
34	36.6	61.0	693	14	AB034324	AB034324 Human imm
35	36.6	61.0	693	14	AB034325	AB034325 Human imm
36	36.6	61.0	693	14	AB034326	AB034326 Human imm
37	36.6	61.0	693	14	AB034327	AB034327 Human imm
38	36.6	61.0	8975	14	D88822	D88822 HIV 1 isola
39	36.6	61.0	9034	14	AF286234	AF286234 HIV 1 str
40	36.6	61.0	9152	14	AF443109	AF443109 HIV 1 iso
41	36.6	61.0	9811	14	CIVCG	X52154 Chimpanzee
42	36	60.0	663	14	HIM274572	AJ274572 Human imm
43	36	60.0	1015	14	HIV1B7163R	L22086 Human immun
44	36	60.0	8853	14	AF361875	AF361875 HIV 1 iso
45	36	60.0	9012	14	AF286235	AF286235 HIV 1 str

ALIGNMENTS

RESULT 1
AX188558
LOCUS AX188558 1608 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 27 from Patent WO0147955.
ACCESSION AX188558
VERSION AX188558.1 GI:15142199
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 1608)
AUTHORS Hanke,T.M. and Mcmichael,A.J.
TITLE Improvements in or relating to immune responses to hiv
JOURNAL Patent: WO 0147955-A 27 05-JUL-2001;
MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine Initiative (US) ; University of Nairobi (KE)
FEATURES Location/Qualifiers
source 1..1608
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Chimeric polynucleotide"
BASE COUNT 314 a 598 c 458 g 238 t
ORIGIN

```
  Query Match             100.0%;  Score 60;  DB 6;  Length 1608;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-07;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaqqgccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     472 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 531


RESULT   2
AX188560
LOCUS       AX188560                1914 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 29 from Patent WO0147955.
ACCESSION   AX188560
VERSION     AX188560.1  GI:15142200
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 1914)
  AUTHORS   Hanke,T.M. and Mcmichael,A.J.
  TITLE     Improvements in or relating to immune responses to hiv
  JOURNAL   Patent: WO 0147955-A 29 05-JUL 2001;
            MEDICAL RESEARCH COUNCIL (GB) , International Aids Vaccine
            Initiative (US) , University of Nairobi (KE)
FEATURES             Location/Qualifiers
     source          1. .1914
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note "Chimeric polynucleotide"
BASE COUNT      383 a    728 c    534 g    269 t
ORIGIN


  Query Match             100.0%;  Score 60;  DB 6;  Length 1914;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-07;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaqqgccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525


RESULT   3
AX188562
LOCUS       AX188562                2493 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 31 from Patent WO0147955.
ACCESSION   AX188562
VERSION     AX188562.1  GI:15142201
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2493)
  AUTHORS   Hanke,T.M. and Mcmichael,A.J.
  TITLE     Improvements in or relating to immune responses to hiv
  JOURNAL   Patent: WO 0147955-A 31 05-JUL-2001;
            MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
            Initiative (US) ; University of Nairobi (KE)
FEATURES             Location/Qualifiers
     source          1. .2493
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Chimeric polynucleotide"
BASE COUNT      516 a    887 c    685 g    405 t
ORIGIN


  Query Match             100.0%;  Score 60;  DB 6;  Length 2493;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-07;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaqqgccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525


RESULT   4
AX188564
LOCUS       AX188564                4350 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 33 from Patent WO0147955.
ACCESSION   AX188564
VERSION     AX188564.1  GI:15142202
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4350)
  AUTHORS   Hanke,T.M. and Mcmichael,A.J.
  TITLE     Improvements in or relating to immune responses to hiv
  JOURNAL   Patent: WO 0147955-A 33 05-JUL-2001;
            MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
            Initiative (US) ; University of Nairobi (KE)
FEATURES             Location/Qualifiers
     source          1. .4350
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note "Chimeric polynucleotide"
BASE COUNT      888 a   1630 c   1259 g    573 t
ORIGIN


  Query Match             100.0%;  Score 60;  DB 6;  Length 4350;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-07;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaqqgccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525


RESULT   5
AF201927
LOCUS       AF201927                1509 bp    DNA     linear   SYN 16-MAR-2000
DEFINITION  Synthetic construct gag protein gene, complete cds.
ACCESSION   AF201927
VERSION     AF201927.1  GI:7248702
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 1509)
  AUTHORS   zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
            Selby,M., Otten,G.R. and Barnett,S.W.
  TITLE     Increased expression and immunogenicity of sequence-modified human
            immunodeficiency virus type 1 gag gene
  JOURNAL   J. Virol. 74 (6), 2628-2635 (2000)
  MEDLINE   20148954
   PUBMED   10684277
REFERENCE   2  (bases 1 to 1509)
  AUTHORS   zur Megede,J. and Barnett,S.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV 1999) Vaccines, Chiron Corporation, 4560 Horton,
            Emeryville, CA 94608, USA
FEATURES             Location/Qualifiers
     source          1. .1509
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             1. .1509
                     /note="HIV-1SF2 p55gag"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id "AAF43628.1"
```

/db_xref "GI:7229426"
/translation="MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERF
AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
KIEEEQNKSKKKAQQAAAAMSNSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKV
VEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVH
PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
RMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKT
ILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNPATIMMQRGNFRNQRKTV
KCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSYKGRPGNF
LQSRPEPTAPPEESFRFGEETTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
BASE COUNT 328 a 543 c 479 g 159 t
ORIGIN

Query Match 97.4%; Score 58.4; DB 12; Length 1509;
Best Local Similarity 98.4%; Pred. No. 2.1e-06;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 856 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 915

RESULT 6
AF202464
LOCUS AF202464 1847 bp DNA linear SYN 12-MAR-2000
DEFINITION Synthetic construct gag protease fusion protein GP1 gene, complete cds.
ACCESSION AF202464
VERSION AF202464.1 GI:7229424
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 1847)
AUTHORS zur Megede,J., Chen,M.C., Lee,B., Schaefer,M., Greer,C.E., Selby,M., Otten,G.R. and Barnett,S.W.
TITLE Increased expression and immunogenicity of sequence modified human immunodeficiency virus type 1 gag gene
JOURNAL J. Virol. 74 (6), 2628-2635 (2000)
MEDLINE 20148954
PUBMED 10684277
REFERENCE 2 (bases 1 to 1847)
AUTHORS zur Megede,J. and Barnett,S.W.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608, USA
FEATURES Location/Qualifiers
source 1..1847
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS 1..1509
/codon_start=1
/transl_table=11
/product="gag protease fusion protein GP1"
/protein_id="AAF42819.1"
/db_xref="GI:7229425"
/translation="MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERF
AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
KIEEEQNKSKKKAQQAAAAMSNSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKV
VEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVH
PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
RMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKT
ILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNPATIMMQRGNFRNQRKTV
KCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSYKGRPGNF
LQSRPEPTAPPEESFRFGEETTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
BASE COUNT 456 a 576 c 665 g 250 t
ORIGIN

Query Match 97.4%; Score 58.4; DB 12; Length 1847;
Best Local Similarity 98.4%; Pred. No. 2e-06;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 856 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 915

RESULT 7
AF202465
LOCUS AF202465 1847 bp DNA linear SYN 12-MAR-2000
DEFINITION Synthetic construct gag protease fusion protein GP2 gene, complete cds.
ACCESSION AF202465
VERSION AF202465.1 GI:7229426
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 1847)
AUTHORS zur Megede,J., Chen,M.C., Lee,B., Schaefer,M., Greer,C.E., Selby,M., Otten,G.R. and Barnett,S.W.
TITLE Increased expression and immunogenicity of sequence modified human immunodeficiency virus type 1 gag gene
JOURNAL J. Virol. 74 (6), 2628-2635 (2000)
MEDLINE 20148954
PUBMED 10684277
REFERENCE 2 (bases 1 to 1847)
AUTHORS zur Megede,J. and Barnett,S.W.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608, USA
FEATURES Location/Qualifiers
source 1..1847
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS 1..1509
/codon_start=1
/transl_table=11
/product="gag protease fusion protein GP2"
/protein_id="AAF42820.1"
/db_xref="GI:7229427"
/translation="MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERF
AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
KIEEEQNKSKKKAQQAAAAMSNSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKV
VEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVH
PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
RMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKT
ILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNPATIMMQRGNFRNQRKTV
KCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSYKGRPGNF
LQSRPEPTAPPEESFRFGEETTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
BASE COUNT 456 a 576 c 665 g 250 t
ORIGIN

Query Match 97.4%; Score 58.4; DB 12; Length 1847;
Best Local Similarity 98.4%; Pred. No. 2e-06;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 856 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 915

RESULT 8
AF287354
LOCUS AF287354 1548 bp DNA linear SYN 01-NOV-2000
DEFINITION Synthetic construct HIV-1 derived gag protein gene, complete cds.
ACCESSION AF287354
VERSION AF287354.1 GI:11066866
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 1548)

```
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Helper plasmids for production of HIV derived vectors
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1548)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2000) Chemical Pathology, Women's and Children's
            Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES             Location/Qualifiers
     source          1..1548
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="derived from Human immunodeficiency virus type 1
                     isolate YU-2"
     CDS             37..1539
                     /note="codon optimized reading frame"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAG28738.1"
                     /db_xref="GI:11066867"
                     /translation="MGARASVLSAGELDKWEKIRLRPGGKKQYRLKHIVWASRELERF
                     AVDPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNIVATLYCVHQKIEVKDTKEALE
                     KIEEEQNKSKKKAQQAAADTGNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVVE
                     EKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPV
                     HAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM
                     YSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTIL
                     KALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKC
                     FNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQ
                     SRPEPTAPSEESVRFGEETTTPSQKQEPIDKELYPLASLRSLFGSDPSSQ"
BASE COUNT      374 a    495 c    515 g    164 t
ORIGIN


  Query Match             81.3%;  Score 48.8;  DB 12;  Length 1548;
  Best Local Similarity   88.3%;  Pred. No. 0.00067;
  Matches   53;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 gacatccggcaggggcccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                |||||| | |||||||||||||||||||||| | |||||||||||| | |||| |||||||
Db        886 GACATCAGGCAGGGCCCCAAGGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 945


RESULT 9
AF287352
LOCUS       AF287352                4352 bp    DNA     linear   SYN 01-NOV-2000
DEFINITION  Synthetic construct HIV-1 derived gag protein and pol protein
            genes, complete cds
ACCESSION   AF287352
VERSION     AF287352.1  GI:11066861
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4352)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Helper plasmids for production of HIV derived vectors
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4352)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2000) Chemical Pathology, Women's and Children's
            Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES             Location/Qualifiers
     source          1..4352
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="derived from Human immunodeficiency virus type 1
                     isolate YU-2"
     CDS             37..1539
                     /note="codon optimized reading frame"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAG28735.1"
                     /db_xref="GI:11066862"
                     /translation="MGARASVLSAGELDKWEKIRLRPGGKKQYRLKHIVWASRELERF
                     AVDPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQKIEVKDTKEALE
                     KIEEEQNKSKKKAQQAAADTGNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVVE
                     EKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPV
                     HAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM
                     YSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTIL
                     KALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKC
                     FNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQ
                     SRPEPTAPSEESVRFGEETTTPSQKQEPIDKELYPLASLRSLFGSDPSSQ"
     CDS             <1338..4343
                     /note="codon optimized reading frame"
                     /codon_start=1
                     /transl_table=11
                     /product="pol protein"
                     /protein_id="AAG28736.1"
                     /db_xref="GI:11066863"
                     /translation="FFREDLAFLQGKAREFSSEQTRANSPTRRELQVWGRDNNSLSEAGA
                     DRQGTVSFSFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMNLPGRWKPKMIG
                     GIGGFIKVRQYDQIPIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIET
                     VPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKK
                     KDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLHED
                     FRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDLVIY
                     QYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKW
                     TVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVIPLTE
                     EAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYA
                     RTRGAHTNDVKQLTEAVQKIATESIVIWGKTPKFKLPIQKETWEIWWTEYWQATWIPE
                     WEFVNTPPLVKLWYQLEKEPIIGAETFYVDGAANRETKLGKAGYVTNKGRQKVVSLTD
                     TTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDRSESELVSQIIEQLIKKE
                     KVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQEEHEKYHSNWRAMASDFN
                     LPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIE
                     AEVIPAETGQETAYFLLKLAGRWPVTTIHTDNGSNFTSATVKAACWWAGIKQEFGIPY
                     NPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIV
                     DIIATDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDNSDIKV
                     VPRRKAKIIRDYGKQMAGDDCVAGRQDED"
BASE COUNT     1066 a   1344 c   1410 g    532 t
ORIGIN


  Query Match             81.3%;  Score 48.8;  DB 12;  Length 4352;
  Best Local Similarity   88.3%;  Pred. No. 0.00073;
  Matches   53;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 gacatccggcaggggcccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
                |||||| | |||||||||||||||||||||| | |||||||||||| | |||| |||||||
Db        886 GACATCAGGCAGGGCCCCAAGGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 945


RESULT 10
AF287353
LOCUS       AF287353                4353 bp    DNA     linear   SYN 01-NOV-2000
DEFINITION  Synthetic construct HIV-1-derived gag-pol fusion protein gene,
            complete cds.
ACCESSION   AF287353
VERSION     AF287353.1  GI:11066864
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4353)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Helper plasmids for production of HIV derived vectors
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4353)
  AUTHORS   Fuller,M. and Anson,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2000) Chemical Pathology, Women's and Children's
            Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES             Location/Qualifiers
     source          1..4353
                     /organism="synthetic construct"
```

```
                     /db_xref "taxon:32630"
                     /note "derived from Human immunodeficiency virus type 1
                     isolate YU-2"
     CDS             37..4344
                     /note "codon optimized reading frame"
                     /codon_start 1
                     /transl_table 11
                     /product "gag pol fusion protein"
                     /protein_id "AAG28737.1"
                     /db_xref "GI:11066865"
                     /translation "M[illegible]VWA[illegible]
                     AV[illegible]SLYNTVA[illegible]
                     [illegible]NAWVKVVE
                     EKAFSPEVIPMFSALSEGA[illegible]AAEWDR[illegible]HPV
                     HAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM
                     YSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTIL
                     KALGPAATLEEMMTACQGVGGPGHKARVLAEAM[illegible]NFRNQRKIVKC
                     FNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGK[illegible]
                     EQIRANSPTRRELQVWRRDNNSLSEAGADRQGTVSFSFPQITLWQRPLVTIKIGGQLK
                     EALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQI[illegible]HKAIGTVLVG
                     PTPVNIIGRNLLTQIGC[illegible]TLNF[illegible]
                     CIEMEK[illegible]
                     HPAGLKKKKSVTVLDVGDAYF[illegible]NVLPQGW
                     KGSPAIFQSSMTKILEPFRKQNP[illegible]
                     WGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQ
                     IYAGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIA
                     EIQKQGQGQWTYQIYQEPFKNLKTGKYARTRGAHTNDVKQLTEAVQKIATESIVIWGK
                     TPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVD
                     GAANRETKLGKAGYVTNKGRQKVVSL[illegible]
                     A[illegible]
                     LFLDGIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCS
                     PGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHT
                     DNGSNFTSATVKAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLK
                     TAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSR
                     DPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED"
BASE COUNT     1051 a   1358 c   1435 g    509 t
ORIGIN

  Query Match             81.3%;  Score 48.8;  DB 12;  Length 4353;
  Best Local Similarity   88.3%;  Pred. No. 0.00073;
  Matches   53;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 gacatccgccaggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||| | |||||||||||||||||||| | |||||||||||| | |||| |||||||
Db        886 GACATCAGGCAGGGCCCCAAGGAGCCCTTCAGGGACTACGTGGACAGGTTCTACAAGACC 945


RESULT 11
AX019132
LOCUS       AX019132                4307 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 2 from Patent WO9941397.
ACCESSION   AX019132
VERSION     AX019132.1  GI:10043165
KEYWORDS    
SOURCE      Human immunodeficiency virus.
  ORGANISM  Human immunodeficiency virus
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 4307)
  AUTHORS   Mitrophanous,K., Kingsman,A.J. and Kim,N.
  TITLE     Anti-viral vectors
  JOURNAL   Patent: WO 9941397-A 2 19-AUG-1999;
            MITROPHANOUS KYRIACOS (GB); KINGSMAN ALAN JOHN (GB); OXFORD
            BIOMEDICA LTD (GB); KIM NARRY (KR)
FEATURES             Location/Qualifiers
     source          1..4307
                     /organism "Human immunodeficiency virus"
                     /db_xref "taxon:12721"
BASE COUNT     1114 a   1214 c   1294 g    685 t
ORIGIN

  Query Match             79.7%;  Score 47.8;  DB 6;  Length 4307;
  Best Local Similarity   88.1%;  Pred. No. 0.0014;
  Matches   52;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 gacatccgccaggccccaaggagcccttccgcgactacgtggaccgcttcttcaagac 59
              |||||||||| ||||| ||||| ||||| |||||||||||||||| |||| ||| ||
Db        850 GACATCGGCCAAGGCCGAAGGAACCCTTCGCGACTACGTGGACCGGTTCTACAAAAC 908


RESULT 12
AX035453
LOCUS       AX035453                4307 bp    DNA     linear   PAT 13-NOV-2000
DEFINITION  Sequence 2 from Patent WO0055341.
ACCESSION   AX035453
VERSION     AX035453.1  GI:11191095
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4307)
  AUTHORS   Uden,M. and Mitrophanous,K.
  TITLE     Anti-viral vectors
  JOURNAL   Patent: WO 0055341-A 2 21-SEP-2000;
            UDEN MARK (GB); OXFORD BIOMEDICA LTD (GB); MITROPHANOUS KYRIACOS
            (US)
FEATURES             Location/Qualifiers
     source          1..4307
                     /organism "synthetic construct"
                     /db_xref "taxon:32630"
                     /note "gagpol-SYNgp codon optimised gagpol sequence"
BASE COUNT     1135 a   1198 c   1285 g    689 t
ORIGIN

  Query Match             79.7%;  Score 47.8;  DB 6;  Length 4307;
  Best Local Similarity   88.1%;  Pred. No. 0.0014;
  Matches   52;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 gacatccgccaggccccaaggagcccttccgcgactacgtggaccgcttcttcaagac 59
              |||||||||| ||||| ||||| ||||| |||||||||||||||| |||| ||| ||
Db        850 GACATCGGCCAAGGCCGAAGGAACCCTTCGCGACTACGTGGACCGGTTCTACAAAAC 908


RESULT 13
AX056836
LOCUS       AX056836                4307 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 2 from Patent WO0075370.
ACCESSION   AX056836
VERSION     AX056836.1  GI:12309774
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4307)
  AUTHORS   Mitrophanous,K., Kim,N.H. and Kotsopoulou,E.
  TITLE     In vivo selection method for determining inhibitory rna molecules
  JOURNAL   Patent: WO 0075370-A 2 14-DEC-2000;
            Oxford Biomedica (UK) Limited (GB)
FEATURES             Location/Qualifiers
     source          1..4307
                     /organism "synthetic construct"
                     /db_xref "taxon:32630"
                     /note "Codon optimised gagpol sequence"
BASE COUNT     1135 a   1198 c   1285 g    689 t
ORIGIN

  Query Match             79.7%;  Score 47.8;  DB 6;  Length 4307;
  Best Local Similarity   88.1%;  Pred. No. 0.0014;
  Matches   52;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 gacatccgccaggccccaaggagcccttccgcgactacgtggaccgcttcttcaagac 59
              |||||||||| ||||| ||||| ||||| |||||||||||||||| |||| ||| ||
```

```
Db    850 GACATCCGCCAAGGCCCGAAGGAACCCTTTCGCGACTACGTGGACCGGTTCTACAAAAC 908


RESULT 14
AX283597
LOCUS       AX283597                4307 bp    DNA     linear   PAT 20-NOV-2001
DEFINITION  Sequence 15 from Patent WO0179518.
ACCESSION   AX283597
VERSION     AX283597.1  GI:17044324
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (sites)
  AUTHORS   Kingsman,A.J., Kim,N.T., Kotsopoulou,E.E., Rohll,J. and
            Mitrophanous,K.A.
  TITLE     Method
  JOURNAL   Patent: WO 0179518-A 15 25-OCT-2001;
            Oxford Biomedica (UK) Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .4307
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Codon optimised HIV gag-pol"
BASE COUNT     1137 a   1192 c   1286 g    692 t
ORIGIN


  Query Match             79.7%;  Score 47.8;  DB 6;  Length 4307;
  Best Local Similarity   88.1%;  Pred. No. 0.0014;
  Matches   52;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 gacatccgccaaggcccgaaggaacccttcgcgactacgtggaccggttctacaaaac 59
          ||||||||||| ||||| ||||| ||||| |||||||||||||||| |||| ||| ||
Db    850 GACATCCGCCAAGGCCCGAAGGAACCCTTTCGCGACTACGTGGACCGGTTCTACAAAAC 908


RESULT 15
AX283607
LOCUS       AX283607                4307 bp    DNA     linear   PAT 20-NOV-2001
DEFINITION  Sequence 25 from Patent WO0179518.
ACCESSION   AX283607
VERSION     AX283607.1  GI:17044333
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (sites)
  AUTHORS   Kingsman,A.J., Kim,N.T., Kotsopoulou,E.E., Rohll,J. and
            Mitrophanous,K.A.
  TITLE     Method
  JOURNAL   Patent: WO 0179518-A 25 25-OCT-2001;
            Oxford Biomedica (UK) Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .4307
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Codon optimised gag-pol sequence (pSYNGP)"
BASE COUNT     1137 a   1192 c   1286 g    692 t
ORIGIN


  Query Match             79.7%;  Score 47.8;  DB 6;  Length 4307;
  Best Local Similarity   88.1%;  Pred. No. 0.0014;
  Matches   52;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 gacatccgccaaggcccgaaggaacccttcgcgactacgtggaccggttctacaaaac 59
          ||||||||||| ||||| ||||| ||||| |||||||||||||||| |||| ||| ||
Db    850 GACATCCGCCAAGGCCCGAAGGAACCCTTTCGCGACTACGTGGACCGGTTCTACAAAAC 908


Search completed: July  2, 2002, 21:57:36

Job time: 8953 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

```
Run on:            July  2, 2002, 21:25:35 ; Search time 2350.28 Seconds
                                              (without alignments)
                                              344.562 Million cell updates/sec

Title:             US-09-475-704-2
Perfect score:     60
Sequence:          1 gacatccgccaggccccaa..........tggaccgcttcttcaagacc 60

Scoring table:     IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0

Searched:          13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters:       27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :         EST:*
                   1:  em_estba:*
                   2:  em_esthum:*
                   3:  em_estin:*
                   4:  em_estmu:*
                   5:  em_estov:*
                   6:  em_estpl:*
                   7:  em_estro:*
                   8:  em_htc:*
                   9:  gb_est1:*
                   10:  gb_est2:*
                   11:  gb_htc:*
                   12:  gb_gss:*
                   13:  em_gss_hum:*
                   14:  em_gss_inv:*
                   15:  em_gss_pln:*
                   16:  em_gss_vrt:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	28.6	47.7	718	10	BG647389	BG647389 EST509008
	2	28.4	47.3	525	10	BG549181	BG549181 947073D08
	3	28.4	47.3	580	10	BG517389	BG517389 947061G10
	4	28.4	47.3	592	10	BG550427	BG550427 947074H04
	5	28.2	47.0	585	9	AW147097	AW147097 7979911F07
	6	28	46.7	524	10	BG313886	BG313886 WHE2066_B
	7	28	46.7	647	10	BF473271	BF473271 WHE0926_E
c	8	27.4	45.7	282	10	BI992820	BI992820 102000000
	9	27.4	45.7	370	10	BE595775	BE595775 PI1_54_G0
	10	27.4	45.7	372	10	BG462293	BG462293 947046F01
	11	27.4	45.7	376	10	BM322819	BM322819 PIC1_10_F
	12	27.4	45.7	398	6	BE364444	BE364444 PI1_14_C0
c	13	27.4	45.7	408	10	BG265186	BG265186 947022F02
	14	27.4	45.7	423	10	BG316623	BG316623 947023E09
c	15	27.4	45.7	449	10	BI319043	BI319043 949038E09
	16	27.4	45.7	455	10	BG559542	BG559542 947077H05
c	17	27.4	45.7	470	10	BG316823	BG316823 947027C02
c	18	27.4	45.7	470	10	BE997330	BE997330 94?017B06
	19	27.4	45.7	473	10	BM324448	BM324448 PIC? ...
c	20	27.4	45.7	478	9	AI973498	AI973498 49? ...
	21	27.4	45.7	480	10	BG549580	BG549580 94?0?2H05
c	22	27.4	45.7	485	10	BG549935	BG549935 947079C12
	23	27.4	45.7	489	10	BG355028	BG355028 94703?F05
	24	27.4	45.7	494	9	AW672420	AW672420 LG1_359_A
c	25	27.4	45.7	502	10	BG264895	BG264895 94702_A11
c	26	27.4	45.7	502	10	BG349909	BG349909 947037?07
	27	27.4	45.7	503	10	BG550414	BG550414 947074F03
	28	27.4	45.7	504	10	BI325390	BI325390 949038D09
c	29	27.4	45.7	505	10	BG517832	BG517832 947071E10
c	30	27.4	45.7	506	10	BG349889	BG349889 94703?G?1
c	31	27.4	45.7	507	10	BG462555	BG462555 947058B05
c	32	27.4	45.7	510	10	BG316900	BG316900 947027H03
	33	27.4	45.7	511	10	BG316576	BG316576 947023H06
c	34	27.4	45.7	512	10	BG462644	BG462644 947058G05
	35	27.4	45.7	514	10	BG317262	BG317262 94702?D01
	36	27.4	45.7	514	10	BI233661	BI233661 949030F05
	37	27.4	45.7	516	10	BG458470	BG458470 947044D09
	38	27.4	45.7	520	9	AW282932	AW282932 LG1_304_D
	39	27.4	45.7	522	10	BE597871	BE597871 PI1_?? ...
	40	27.4	45.7	524	10	BM318585	BM318585 PI1_?? ...
c	41	27.4	45.7	528	10	BG458654	BG458654 947052D12
c	42	27.4	45.7	528	10	BG462610	BG462610 947058E09
c	43	27.4	45.7	537	10	BG316758	BG316758 947026F03
c	44	27.4	45.7	537	10	BG517229	BG517229 947062F1?
	45	27.4	45.7	537	10	BG549980	BG549980 94???9H??

ALIGNMENTS

```
RESULT   1
BG647389
LOCUS       BG647389                 718 bp    mRNA    linear   EST 24-APR-2001
DEFINITION  EST509008 HOGA Medicago truncatula cDNA clone pHOGA-16B18 5' end,
            mRNA sequence.
ACCESSION   BG647389
VERSION     BG647389.1  GI:13782501
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
            Utterback,T., Cho,J. and Fraser,C.M.
  TITLE     ESTs from roots of Medicago truncatula treated with
            oligogalacturonides of DP 6-20
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Michael G. Hahn
            Complex Carbohydrate Research Center
            University of Georgia
            220 Riverbend Road, Athens, GA 30602-4712, USA
            Tel: 706 542 4457
            Fax: 706 542 4412
            Email: hahn@ccrc.uga.edu
            C99096 TIGR sequence name: MTMBC09IK More information is
            available at: www.medicago.org
            Seq primer: SKmod (CTA GAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1..718
                     /organism="Medicago truncatula"
                     /cultivar="A17"
                     /db_xref="taxon:3880"
                     /clone="pHOGA-16B18"
                     /clone_lib="HOGA"
                     /tissue_type="3 day old seedling roots"
                     /dev_stage="24 hours after treatment in the dark at 26 C
                     with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
```

presence of 100 ug/ml Gentamicin"
/lab_host="XLOLR"
/note "Vector: pBluescript SK ; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda Zap phage using Ex assist helper phage and propagated in SOLR cells."
BASE COUNT 218 a 207 c 140 g 153 t
ORIGIN

Query Match 47.7%; Score 28.6; DB 10; Length 713;
Best Local Similarity 67.8%; Pred. No. 1.2e+02;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 2 acatccgccaagqgccccaaggagccctccgcgactacqtqqaccqcttcttcaaqacc 60
|| ||||||| | ||| || ||| |||| | || ||| | | | | ||| ||| ||
Db 584 ACGTCCGCCAAGACCCAAACGAGTCCTTGCCAGAATACCTCGTCTGATTCAACAACGCC 642

RESULT 2
BG549181
LOCUS BG549181 525 bp mRNA linear EST 05-APR-2001
DEFINITION 947073D08.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA sequence.
ACCESSION BG549181
VERSION BG549181.1 GI:13557825
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 525)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 947073 row: D column: 08.
FEATURES Location/Qualifiers
source 1. .525
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="947 - 2 week shoot from Barkan lab"
/tissue_type="leaf and stem, including leaf base"
/dev_stage="2 week old seedling (3 leaves)"
/lab_host="XL1-Blue"
/note="Organ: shoot, Vector: Lambda ZAP (pBlueScript SK-), Site_1: EcoRI; Site_2: XhoI, Directionally cloned using Stratagene's Unizap XR cDNA cloning kit with the 5' end at the EcoRI site. The library represents 8 x 10e5 independent recombinant phage. The plants were greenhouse grown."
BASE COUNT 92 a 171 c 170 g 92 t
ORIGIN

Query Match 47.3%; Score 28.1; DB 10; Length 525;
Best Local Similarity 70.4%; Pred. No. 1.3e+02;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 7 cqccaqqgccccaaggagccctccgcgactacqtqqaccqcttcttcaaqacc 60
|| |||| | |||| | ||| | ||| | || ||||| | |||||||||| |
Db 282 CGGCAGGACATCAAGAACCCCGGCAGCGTCAACCAGGACCCCATCTTCAAGAGC 335

RESULT 3
BG517389
LOCUS BG517389 580 bp mRNA linear EST 30-MAR-2001
DEFINITION 947061D3.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA sequence.
ACCESSION BG517389
VERSION BG517389.1 GI:13490625
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 580)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 947061 row: G column: 10.
FEATURES Location/Qualifiers
source 1. .580
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="947 - 2 week shoot from Barkan lab"
/tissue_type="leaf and stem, including leaf base"
/dev_stage="2 week old seedling (3 leaves)"
/lab_host="XL1-Blue"
/note="Organ: shoot, Vector: Lambda ZAP (pBlueScript SK-), Site_1: EcoRI; Site_2: XhoI, Directionally cloned using Stratagene's Unizap XR cDNA cloning kit with the 5' end at the EcoRI site. The library represents 8 x 10e5 independent recombinant phage. The plants were greenhouse grown."
BASE COUNT 84 a 216 c 178 g 102 t
ORIGIN

Query Match 47.3%; Score 28.4; DB 10; Length 580;
Best Local Similarity 70.4%; Pred. No. 1.3e+02;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 7 cqccaqqgccccaaggagccctccgcgactacqtqqaccqcttcttcaaqacc 60
|| |||| | |||| | ||| | ||| | || ||||| | |||||||||| |
Db 508 CGGCAGGACATCAAGAACCCCGGCAGCGTCAACCAGGACCCCATCTTCAAGAGC 561

RESULT 4
BG550427
LOCUS BG550427 592 bp mRNA linear EST 09-APR-2001
DEFINITION 947074B04.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA sequence.
ACCESSION BG550427
VERSION BG550427.1 GI:13562207
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 592)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

```
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 947074  row: H  column: 04.
FEATURES             Location/Qualifiers
     source          1..592
                     /organism "Zea mays"
                     /cultivar "B73"
                     /db_xref "taxon:4577"
                     /clone_lib "947 - 2 week shoot from Barkan lab"
                     /tissue_type "leaf and stem, including leaf base"
                     /dev_stage "2 week old seedling (3 leaves)"
                     /lab_host "XL1 Blue"
                     /note "Organ: shoot; Vector: Lambda ZAP (pBlueScript SK-);
                     Site_1: EcoRI; Site_2: XhoI, Directionally cloned using
                     Stratagene's UniZap XR cDNA cloning kit with the 5' end
                     at the EcoRI site. The library represents 8 x 10e5
                     independent recombinant phage. The plants were greenhouse
                     grown."
BASE COUNT     87 a    212 c    194 g     99 t
ORIGIN

  Query Match            47.3%;  Score 28.4;  DB 10;  Length 592;
  Best Local Similarity   70.4%;  Pred. No. 1.4e+02;
  Matches   38;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          7 [illegible] 60
Db        468 [illegible] 521


RESULT   5
AW147097
LOCUS       AW147097                 585 bp    mRNA    linear   EST 03-NOV-1999
DEFINITION  707011F07.x1 707 - Mixed adult tissues from Walbot lab (SK) Zea
            mays cDNA, mRNA sequence.
ACCESSION   AW147097
VERSION     AW147097.1  GI:6194993
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 707011  row: F  column: 07.
FEATURES             Location/Qualifiers
     source          1..585
                     /organism "Zea mays"
                     /cultivar "W23"
                     /db_xref "taxon:4577"
                     /clone_lib "707 - Mixed adult tissues from Walbot lab (SK
                     )"
                     /tissue_type "tassel, kernel, silk, husk, root, leaf"
                     /dev_stage "adult"
                     /lab_host "DH10B"
                     /note "organ: tassel, kernel, silk, husk, root, leaf;
                     Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
                     differentiated maize tissues from an active  Mutator
                     plant. Tissue ratio is 4/2/?/?/?/1 (tassel, kernel, silk,
                     husk, root, leaf). Unidirectionally cloned."
BASE COUNT     79 a    221 c    202 g     83 t
ORIGIN

  Query Match            47.0%;  Score 28.2;  DB 9;  Length 585;
  Best Local Similarity   68.4%;  Pred. No. 1.5e+02;
  Matches   39;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 [illegible] 57
Db         39 [illegible] 95


RESULT   6
BG313886
LOCUS       BG313886                 524 bp    mRNA    linear   EST 23-FEB-2001
DEFINITION  WHE2066_B07_D14ZS Wheat salt stressed sheath cDNA library Triticum
            aestivum cDNA clone WHE2066_B07_D14, mRNA sequence.
ACCESSION   BG313886
VERSION     BG313886.1  GI:13115689
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 524)
  AUTHORS   Akhunov,E., Anderson,O.D., Chao,S., Chin,A., Choi,D.W., Close,T.J.,
            Dvorak,J., Fenton,R.D., Han,P.S., Hsia,C.C., Kang,Y., Kianian,F.,
            Lazo,G.R., Miller,R., Otto,C., Nguyen,H.T., Rausch,C.J., Seaton
            ,C.L., Simons,K., Tong,J.C. and Zhang,D.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Salt stressed sheath cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1..524
                     /organism "Triticum aestivum"
                     /cultivar "Chinese Spring"
                     /db_xref "taxon:4565"
                     /clone "WHE2066_B07_D14"
                     /clone_lib "Wheat salt stressed sheath cDNA library"
                     /tissue_type "Sheath"
                     /dev_stage "Adult plant"
                     /lab_host "E. coli SOLR"
                     /note "Vector: Lambda Uni ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown under
                     hydroponic conditions at UC Davis, salt stressed  for 12
                     hours, and for 7 days, then dissected and frozen (Akhunov
                     in J. Dvorak lab). Total RNA was prepared from sheath
                     tissue, equal quantities  of RNA were pooled from the two
                     samples, polyA was purified from the  pooled RNA, a cDNA
                     library was made, and the cDNA clones were in vivo
                     excised to give pBluescript phagemids in the TJ Close lab
                     at the University of California, Riverside (Akhunov, Chin
                     , Choi, Close, Fenton,  Kianian, Otto, Simons, Zhang).
                     Plasmid DNA preparations and DNA  sequencing were
                     performed in the OD Anderson lab (all other authors)."
```

```
BASE COUNT       86 a    165 c    173 g    100 t
ORIGIN


  Query Match             46.7%;  Score 28;  DB 10;  Length 521;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaggaccccaaggaggccttccgcgactacgtgaaccgcttcttcaaggacc 60
           | | ||||| | | |  |  |||| ||| ||||| |    || |||||||||  ||| |||
Db      55 GGCCTCCCCGACGCCGACCTGGAGGCCTCCCGCGCCGCGCTGCACCGCTTCGGCAACACC 114


RESULT   7
BF473271
LOCUS       BF473271                 647 bp    mRNA    linear   EST 04-DEC-2000
DEFINITION  WHE0926_E11_122Z9 Wheat 5 15 DAP spike cDNA library Triticum
            aestivum cDNA clone WHE0926_E11_122, mRNA sequence.
ACCESSION   BF473271
VERSION     BF473271.1  GI:11542453
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - 5-15 DAP spike cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism "Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref "taxon:4565"
                     /clone "WHE0926_E11_122"
                     /clone_lib "Wheat 5 15 DAP spike cDNA library"
                     /tissue_type="Spike"
                     /dev_stage="Adult plant"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                     greenhouse. Spikes at 5, 10 and 15 DAP were harvested,
                     total RNA and poly(A) RNA were prepared, a cDNA library
                     was made, and the cDNA clones were in vivo excised to
                     give pBluescript phagemids in the TJ Close lab (Choi,
                     Close, Fenton) at the University of California,
                     Riverside. Plasmid DNA preparations and DNA sequencing
                     were performed in the OD Anderson lab (all other authors
                     )."
BASE COUNT      132 a    184 c    191 g    140 t
ORIGIN


  Query Match             46.7%;  Score 28;  DB 10;  Length 647;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+02;
  Matches   40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaggaccccaaggaggccttccgcgactacgtgaaccgcttcttcaaggacc 60
           | | ||||| | | |  |  |||| ||| ||||| |    || |||||||||  ||| |||
Db      58 GGCCTCCCCGACGCCGACCTGGAGGCCTCCCGCGCCGCGCTGCACCGCTTCGGCAACACC 117


RESULT   8
BI992820
LOCUS       BI992820                 292 bp    mRNA    linear   EST 24-OCT-2001
DEFINITION  1020068G03 yj 1020 - Unigene 1T from Maize Genome Project Zea mays
            cDNA, mRNA sequence.
ACCESSION   BI992820
VERSION     BI992820.1  GI:16479210
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 282)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 724 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 1020068  row: G  column: 03.
FEATURES             Location/Qualifiers
     source          1. .282
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
                     /clone_lib "1020 - Unigene 1T from Maize Genome Project"
                     /note "This library represents the unique genes found in
                     the second round of EST sequencing at Stanford University
                     for the maize genome project. Sequences are present from
                     libraries 947 and 949. Contigs were assembled using
                     ZmDBAssembler and 2 representatives from each contig were
                     selected for the Unigene set. All singlets were also
                     selected."
BASE COUNT       40 a     75 c    104 g     63 t
ORIGIN


  Query Match             45.7%;  Score 27.4;  DB 10;  Length 282;
  Best Local Similarity   69.8%;  Pred. No. 2.3e+02;
  Matches   37;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       8 gccaggaccccaaggaggccttccgcgactacgtgaaccgcttcttcaaggacc 60
           | |||| |  |||| | |||  | ||| | ||  ||||| | |||||||||| |
Db     278 GGCACGACATCAAGAACCCCGGCAGCGTCAACCAGGACCCCATCTTCAAGAGC 226


RESULT   9
BE595775
LOCUS       BE595775                 370 bp    mRNA    linear   EST 18-AUG-2000
DEFINITION  PI1_54_C06.g1_A002 Pathogen induced 1 (PI1) Sorghum bicolor cDNA,
            mRNA sequence.
ACCESSION   BE595775
VERSION     BE595775.1  GI:9850848
KEYWORDS    EST.
SOURCE      Sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: pathogen induced plants
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier Pratt MM
```

```
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602 7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 112
            High quality sequence stop: 429
            POLYA Yes.
FEATURES             Location/Qualifiers
     source          1..379
                     /organism "Sorghum bicolor"
                     /db_xref "taxon:4558"
                     /clone_lib "Pathogen induced 1 (PI1)"
                     /note "Organ: Anthracnose infected leaves from
                     two-week-old sorghum plants 48 hr after inoculation;
                     Vector: pBluescript II from Lambda Zap II; Site_1: XhoI;
                     Site_2: EcoRI; Two week old sorghum plants (BTX 623
                     cultivar) were infected with pathogen (isolate FRM421 of
                     Colletotrichum graminicola, which is a sorghum isolate).
                     RNA was prepared from infected leaves harvested from 45
                     seedlings  48 hours after inoculation.  Note: young
                     seedlings (2 weeks old) exhibit juvenile resistant
                     reaction, which is an incompatible interaction. As they
                     grow older (4 weeks or older), plants resume susceptiblity
                     to anthracnose disease. The library was made from poly-A
                     RNA in the cloning vector lambda ZAP II. Clones to be
                     sequenced were prepared by mass excision. WARNING: While
                     most or all ESTs are expected to derive from the host
                     plant, no effort was made to eliminate ESTs deriving from
                     the pathogen."
BASE COUNT       85 a    120 c     96 g     69 t
ORIGIN


  Query Match             45.7%;  Score 27.4;  DB 10;  Length 370;
  Best Local Similarity   69.8%;  Pred. No. 2.4e+02;
  Matches   37;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          8 gccaggacccaaggagccttccgcgactacgtggaccgcttcttcaagacc 60
               | |||| |  |||| | |||   | ||| | ||  ||||| | ||||||||| |
Db         51 GGCAGGACATCAAGAACCCGGCAGCGTCAACCAGGACCCATCTTCAAGAGC 103


RESULT 10
BG462293
LOCUS       BG462293                 372 bp    mRNA    linear   EST 28 MAR 2001
DEFINITION  947046F01.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
            sequence.
ACCESSION   BG462293
VERSION     BG462293.1  GI:13388362
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 372)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 947046  row: F  column: 01.
FEATURES             Location/Qualifiers
     source          1..372
                     /organism "Zea mays"
                     /cultivar "B73"
                     /db_xref "taxon:4577"
                     /clone_lib "947 - 2 week shoot from Barkan lab"
                     /tissue_type "leaf and stem, including leaf base"
                     /dev_stage "2 week old seedling (3 leaves)"
                     /lab_host "XL1-Blue"
                     /note "Organ: shoot; Vector: Lambda ZAP (pBlueScript SK-);
                     Site_1: EcoRI; Site_2: XhoI; Directionally cloned using
                     Stratagene's UniZap XR cDNA cloning kit with  the 5' end
                     at the EcoRI site. The library represents 8 x 10e5
                     independent recombinant phage. The plants were greenhouse
                     grown."
BASE COUNT       66 a    111 c    133 g     62 t
ORIGIN


  Query Match             45.7%;  Score 27.4;  DB 10;  Length 372;
  Best Local Similarity   69.8%;  Pred. No. 2.4e+02;
  Matches   37;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          8 gccaggacccaaggagccttccgcgactacgtggaccgcttcttcaagacc 60
               | |||| |  |||| | |||   | ||| | ||  ||||| | ||||||||| |
Db        272 GGCAGGACATCAAGAACCCGGCAGCGTCAACCAGGACCCATCTTCAAGAGC 324


RESULT 11
BM322819
LOCUS       BM322819                 376 bp    mRNA    linear   EST 04 JAN 2002
DEFINITION  PIC1_10_F02.b1 A002 Pathogen infected compatible 1 (PIC1) Sorghum
            bicolor cDNA, mRNA sequence.
ACCESSION   BM322819
VERSION     BM322819.1  GI:18060080
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R.,
            Sudman,M. and Pratt,L.H.
  TITLE     An EST database from Sorghum: plants infected with a compatible
            pathogen
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602 7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector, and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20. Three prime sequences, which are obtained with PolyTMix or
            T7 sequencing primer, are presented as the reverse complement.
            Seq primer: JEN REV
            High quality sequence stop: 374
            POLYA No.
FEATURES             Location/Qualifiers
     source          1..376
                     /organism "Sorghum bicolor"
                     /cultivar "BTx623"
                     /db_xref "taxon:4558"
                     /clone_lib "Pathogen infected compatible 1 (PIC1)"
                     /tissue_type "leaves"
                     /dev_stage "4-week old seedlings infected with
                     Colletotrichum graminicola"
                     /note "Vector: pBluescript II SK( ) from Lambda Zap II;
                     Site_1: XhoI; Site_2: EcoRI; Four week old sorghum
```

```
                     seedlings were sprayed with spore suspension prepared from
                     3-week old FRM421, a sorghum isolate of the anthracnose
                     pathogen Colletotrichum graminicola. Inoculated plants
                     were kept in a 25 C dark growth chamber with 100% relative
                     humidity for 24 hr, followed by 12/12 hr of light/dark
                     cycle at 25 C with 90% relative humidity for another 24
                     hr. All leaves were harvested and quick frozen with liquid
                     nitrogen and stored in a -80 C freezer. The library was
                     made from poly A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision.
                     WARNING: While most or all ESTs are expected to derive
                     from the host plant, no effort was made to eliminate ESTs
                     deriving from the pathogen."
BASE COUNT       65 a    122 c    128 g     61 t
ORIGIN


  Query Match             45.7%;  Score 27.4;  DB 10;  Length 476;
  Best Local Similarity   69.8%;  Pred. No. 2.4e+02;
  Matches   37;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       8 qccaqqqccccaaqqaaccctteeqeqactacqtqqaccgcttcttcaaqacc 60
           | |||| |  |||| | |||  | ||| | ||  ||||| | |||||||||| |
Db     250 GGCAGCACATCAACAACCCCGGCAGCGTCAACCAGCACCCCATCTTCAAGAGC 302


RESULT  12
BE364444
ID   BE364444   standard; RNA; EST; 398 BP.
XX
AC   BE364444;
XX
SV   BE364444.1
XX
DT   28-JUL-2000 (Rel. 64, Created)
DT   28-JUL-2000 (Rel. 64, Last updated, Version 1)
XX
DE   P11_14_C08.b1_A002 Pathogen induced 1 (PI1) Sorghum bicolor cDNA, mRNA
DE   sequence.
XX
KW   EST.
XX
OS   Sorghum bicolor (sorghum)
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta, Magnoliophyta, Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Sorghum.
XX
RN   [1]
RP   1-398
RA   Cordonnier-Pratt M.-M., Gingle A., Dean R., Sudman M., Pratt L.H.;
RT   "An EST database from Sorghum: pathogen-induced plants";
RL   Unpublished.
XX
CC   Contact: Cordonnier-Pratt MM
CC   Department of Botany
CC   The University of Georgia
CC   Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
CC   Tel: 706 542 1860
CC   Fax: 706 542 1805
CC   Email: mmpratt@uga.edu
CC   Sequences have been trimmed to exclude PolyA, vector and regions
CC   below Phred quality 16. The threshold for highest quality sequence
CC   is 20.
CC   Seq primer: JEN REV
CC   High quality sequence stop: 396
CC   POLYA=No.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .398
FT                   /db_xref="taxon:4558"
FT                   /db_xref "ESTLIB.5441"
FT                   /note "Organ: Anthracnose infected leaves from two week old
FT                   sorghum plants 48 hr after inoculation; Vector: pBluescript
FT                   II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI;
FT                   Two-week-old sorghum plants (BTX 623 cultivar) were
FT                   inoculated with pathogen (isolate FRM421 of Colletotrichum
FT                   graminicola, which is a sorghum isolate). RNA was prepared
FT                   from inoculated leaves harvested from 45 seedlings 48 hours
FT                   after inoculation. Note: young seedlings (2 weeks old)
FT                   exhibit juvenile resistant reaction, which is an
FT                   incompatible interaction. As they grow older (4 weeks or
FT                   older), plants resume susceptibility to anthracnose disease.
FT                   The library was made from poly-A RNA in the cloning vector
FT                   lambda ZAP II. Clones to be sequenced were prepared by mass
FT                   excision. WARNING: While most or all ESTs are expected to
FT                   derive from the host plant, no effort was made to eliminate
FT                   ESTs deriving from the pathogen."
FT                   /organism="Sorghum bicolor"
FT                   /clone_lib "Pathogen induced 1 (PI1)"
XX
SQ   Sequence 398 BP; 68 A; 128 C; 136 G; 66 T; 0 other;


  Query Match             45.7%;  Score 27.4;  DB 6;  Length 398;
  Best Local Similarity   69.8%;  Pred. No. 2.5e+02;
  Matches   37;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       8 qccaqqqccccaaqqaaccctteeqeqactacqtqqaccgcttcttcaaqacc 60
           | |||| |  |||| | |||  | ||| | ||  ||||| | |||||||||| |
Db     271 GGCAGGACATCAAGAACCCCGGCAGCGTCAACCAGGACCCCATCTTCAAGAGC 323


RESULT  13
BG265180/c
LOCUS       BG265180                 408 bp    mRNA    linear   EST 20-FEB-2001
DEFINITION  947022F02.x2 947   2 week shoot from Parkan lab 2nd ways cDNA, mRNA
            sequence.
ACCESSION   BG265180
VERSION     BG265180.1  GI:12908234
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 947022  row: F  column: 02.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Zea mays"
                     /cultivar "B73"
                     /db_xref "taxon:4577"
                     /clone_lib "947   2 week shoot from Parkan lab"
                     /tissue_type "leaf and stem, including leaf base"
                     /dev_stage="2 week old seedling (3 leaves)"
                     /lab_host "XL1-Blue"
                     /note "Organ: shoot; Vector: Lambda ZAP (pBluescript SK ):
                     Site_1: EcoRI; Site_2: XhoI; Directionally cloned using
                     Stratagene's UniZap XR cDNA cloning kit with the 5' end
                     at the EcoRI site. The library represents 8 x 10e5
                     independent recombinant phage. The plants were greenhouse
                     grown."
BASE COUNT       71 a    115 c    131 g     91 t
ORIGIN
```

```
  Query Match              45.7%;  Score 27.4;  DB 10;  Length 408;
  Best Local Similarity   69.8%;  Pred. No. 2.5e+02;
  Matches   37;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          8 gccaggacccaaggagcccttccgcgactacgtggacccgcttcttcaagacc 60
              | ||| |  |||| | |||  | |||  . ||  ||||| | |||||||| |
Db        405 GGCAGGACATCAAGAACCCGGCAGCGTCAACCAGGACCCATCTTCAAGAGC 353


RESULT 14
BG316623
LOCUS       BG316623                 420 bp    mRNA    linear   EST 26-FEB-2001
DEFINITION  947021E09.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
            sequence.
ACCESSION   BG316623
VERSION     BG316623.1  GI:13125948
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 420)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 947021  row: E  column: 09.
FEATURES             Location/Qualifiers
     source          1..420
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="947 - 2 week shoot from Barkan lab"
                     /tissue_type="leaf and stem, including leaf base"
                     /dev_stage="2 week old seedling (3 leaves)"
                     /lab_host="XL1 Blue"
                     /note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-);
                     Site_1: EcoRI; Site_2: XhoI; Directionally cloned using
                     Stratagene's UniZap XR cDNA cloning kit with  the 5' end
                     at the EcoRI site. The library represents 8 x 10e5
                     independent recombinant phage. The plants were greenhouse
                     grown."
BASE COUNT       70 a    136 c    142 g     72 t
ORIGIN


  Query Match              45.7%;  Score 27.4;  DB 10;  Length 420;
  Best Local Similarity   69.8%;  Pred. No. 2.5e+02;
  Matches   37;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          8 gccaggacccaaggagcccttccgcgactacgtggacccgcttcttcaagacc 60
              | |||  |  |||| . .||  | ||| | ||  ||||| | |||||||| |
Db        408 GGCAGGACATCAAGAACCCGGCAGCGTCAACCAGGACCCATCTTCAAGAGC 368


RESULT 15
BI319043/c
LOCUS       BI319043                 449 bp    mRNA    linear   EST 23-JUL-2001
DEFINITION  949038D09.x2 949 - Juvenile leaf and shoot cDNA from Steve Moose
            Zea mays cDNA, mRNA sequence.
ACCESSION   BI319043
VERSION     BI319043.1  GI:14996878
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 949038  row: D  column: 09.
FEATURES             Location/Qualifiers
     source          1..449
                     /organism="Zea mays"
                     /cultivar="W64A"
                     /db_xref="taxon:4577"
                     /clone_lib="949 - Juvenile leaf and shoot cDNA from Steve
                     Moose"
                     /tissue_type="immature leaf primordium and vegetative
                     meristem"
                     /dev_stage="4 stages from 3-13 days after imbibing"
                     /lab_host="E. coli XLOLR"
                     /note="Organ: juvenile vegetative shoots; Vector:
                     pAD-GAL4 2.1; Site_1: EcoRI; Site_2: XhoI; Equal amounts
                     of total RNA by weight from 4 tissue sources (see below)
                     were pooled, polyA+ RNA isolated, and cDNA synthesized for
                     EcoRI (5') and XhoI (3') directional cloning into lambda
                     Hybrizap vector from Stratagene. Tissue Sources: 1. Whole
                     shoots 3 days after sowing/imbibing in wet soil. 2. Basal
                     1.5 cm shoots 6 days after sowing - includes yellow
                     portions of developing leaves 1-5, primordia from 6-8, and
                     the vegetative apex. 3. Non-green portions of developing
                     leaves 4-5 and the vegetative apex, - including adult leaf
                     primordia, 9 days after sowing. 4. Partially expanded and
                     greening leaves 4-5 at 13 days after sowing."
BASE COUNT       76 a    130 c    153 g     90 t
ORIGIN


  Query Match              4_.7%;  Score 27.4;  DB 10;  Length 449;
  Best Local Similarity   69.8%;  Pred. No. 2.5e+02;
  Matches   37;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          8 gccaggacccaaggagcccttccgcgactacgtggacccgcttcttcaagacc 60
              | |||| |  |||| | |||  | ||| | ||  ||||| | ||||| ||| |
Db        333 GGCAGGACATCAAGAACCCGGCAGCGTCAACCAGGACCCATCTTCAAGAGC 281


Search completed: July  2, 2002, 21:25:38
Job time: 7555 secs
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 21:58:57 ; Search time 63.61 Seconds
(without alignments)
231.693 Million cell updates/sec

Title: US-09-475-704-2
Perfect score: 60
Sequence: 1 gacatccgcccagggcccccaa..........tggaccgcttcttcaagacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: [illegible]
2: [illegible]
3: [illegible]
4: [illegible]
5: [illegible]
6: [illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33.4	55.7	718	6	5204259-2	Patent No. 5204259
2	33.4	55.7	793	6	5204259-4	Patent No. 5204259
3	33.4	55.7	796	6	5204259-6	Patent No. 5204259
4	33.4	55.7	845	1	US-08-589-446-3	Sequence 3, Appli
5	33.4	55.7	845	1	US-08-444-882-3	Sequence 3, Appli
6	33.4	55.7	845	1	US-08-389-459A-3	Sequence 3, Appli
7	33.4	55.7	845	3	US-08-987-867A-3	Sequence 3, Appli
8	33.4	55.7	871	6	5204259-10	Patent No. 5204259
9	33.4	55.7	1021	6	5204259-8	Patent No. 5204259
10	33.4	55.7	5362	3	US-08-463-210-5	Sequence 5, Appli
11	33.4	55.7	7228	2	US-08-850-049-128	Sequence 128, App
12	33.4	55.7	7228	2	US-08-850-049-129	Sequence 129, App
13	33.4	55.7	7228	2	US-08-050-478-128	Sequence 128, App
14	33.4	55.7	7228	2	US-08-050-478-129	Sequence 129, App
15	33.4	55.7	7228	4	US-09-414-117-128	Sequence 128, App
16	33.4	55.7	7228	4	US-09-414-117-129	Sequence 129, App
17	33.4	55.7	7399	2	US-08-418-848A-9	Sequence 9, Appli
18	33.4	55.7	9709	2	US-08-188-583-5	Sequence 5, Appli
19	33.4	55.7	9709	3	US-08-388-353-1	Sequence 1, Appli
20	33.4	55.7	9709	3	US-08-488-551B-1	Sequence 1, Appli
21	33.4	55.7	9739	1	US-08-022-835-1	Sequence 1, Appli
22	33.4	55.7	9739	1	US-08-388-809-1	Sequence 1, Appli
23	33.4	55.7	9739	2	US-08-647-714-1	Sequence 1, Appli
24	33.4	55.7	9746	1	US-08-022-835-3	Sequence 3, Appli
25	33.4	55.7	9746	1	US-08-388-809-3	Sequence 3, Appli
26	33.4	55.7	9746	2	US-08-647-714-3	Sequence 3, Appli
27	33.4	55.7	12494	4	US-08-935-312-13	Sequence 13, Appl
28	33.4	55.7	12494	4	US-08-848-760B-33	Sequence 33, Appl
29	33.4	55.7	15581	3	US-08-646-538-35	Sequence 35, Appl
30	33.4	55.7	15581	4	US-09-503-222-35	Sequence 35, Appl
31	31.8	53.0	1014	1	US-07-648-796A-9	Sequence 9, Appli
32	31.8	53.0	1377	1	US-07-648-796A-13	Sequence 13, Appl
33	31.8	53.0	2313	1	US-07-648-796A-16	Sequence 16, Appl
34	31.8	53.0	2400	1	US-07-648-796A-15	Sequence 15, Appl
35	31.8	53.0	3807	2	US-08-417-210A-78	Sequence 78, Appl
36	31.8	53.0	9207	3	US-08-388-353-800	Sequence 800, App
37	31.8	53.0	9207	3	US-08-488-551B-800	Sequence 800, App
38	31.8	53.0	9737	2	US-08-944-449-7	Sequence 7, Appli
39	30.2	50.3	1314	3	US-08-392-794A-1	Sequence 1, Appli
40	30.2	50.3	2095	1	US-08-333-901-5	Sequence 5, Appli
41	30.2	50.3	2095	1	US-08-456-582-5	Sequence 5, Appli
42	30.2	50.3	2095	2	US-08-898-789-5	Sequence 5, Appli
43	30.2	50.3	8932	4	US-09-124-900-1	Sequence 1, Appli
44	30.2	50.3	8933	3	US-08-463-210-4	Sequence 4, Appli
45	30.2	50.3	8933	4	US-09-620-958A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
5204259-2
;Patent No. 5204259
; APPLICANT: HELLING, TORSTEN B.;CREVIN, HAKAN;NUNN,
;MICHAEL F.
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/344,237
; FILING DATE: 26 APR 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 191,299
; FILING DATE: 06 MAY 1988
; APPLICATION NUMBER: 191,229
; FILING DATE: 06 MAY 1988
; APPLICATION NUMBER: 206,499
; FILING DATE: 13 JUN 1988
; APPLICATION NUMBER: 258,016
; FILING DATE: 14 OCT 1988
;SEQ ID NO:2:
; LENGTH:718
5204259-2

Query Match 55.7%; Score 33.4; DB 6; Length 718;
Best Local Similarity 72.9%; Pred. No. 0.017;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 gacatccgcccagggcccccaaggagcccttccgcgactacgtggaccgcttcttcaagac 59
Db 472 gacataagacaagaccaaggaacccttagagactatgtagaccggttctataaaac 530

RESULT 2
5204259-4
;Patent No. 5204259
; APPLICANT: HELLING, TORSTEN B.;CREVIN, HAKAN;NUNN,
;MICHAEL F.
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/344,237
; FILING DATE: 26 APR 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 191,299
; FILING DATE: 06 MAY 1988
; APPLICATION NUMBER: 191,229
; FILING DATE: 06-MAY-1988

; APPLICATION NUMBER: 206,499
; FILING DATE: 13-JUN-1988
; APPLICATION NUMBER: 258,016
; FILING DATE: 14-OCT-1988
;SEQ ID NO:4:
; LENGTH:793
5204259-4

Query Match 55.7%; Score 33.4; DB 6; Length 793;
Best Local Similarity 72.9%; Pred. No. 0.017;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 gacatccgcccaggcccccaagaagaccctttcagagactatgtagaccggttctataaaaac 59
||||| | || || || ||||| ||||| | ||||| || ||||| |||| || ||
Db 472 gacataagacaaggaccaaaggaacccttcagagactatgtagaccggttctataaaaac 530

RESULT 3
5204259-6
;Patent No. 5204259
; APPLICANT: HEITING,TORSTEN B.;CREVIN, HAKAN;NONN,
;MICHAEL F.
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/344,247
; FILING DATE: 26-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 191,299
; FILING DATE: 06-MAY-1988
; APPLICATION NUMBER.191,229
; FILING DATE: 06-MAY-1988
; APPLICATION NUMBER: 206,499
; FILING DATE: 13-JUN-1988
; APPLICATION NUMBER: 258,016
; FILING DATE: 14-OCT-1988
;SEQ ID NO:6:
; LENGTH:796
5204259-6

Query Match 55.7%; Score 33.4; DB 6; Length 796;
Best Local Similarity 72.9%; Pred. No. 0.017;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 gacatccgcccaggcccccaagaagaccctttcagagactatgtagaccggttctataaaaac 59
||||| | || || || ||||| ||||| | ||||| || ||||| |||| || ||
Db 472 gacataagacaaggaccaaaggaacccttcagagactatgtagaccggttctataaaaac 530

RESULT 4
US-08-589-446-3
; Sequence 3, Application US/08589446
; Patent No. 5614413
; GENERAL INFORMATION:
; APPLICANT: Morrow, Casey D.
; TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC
; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,446
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,009
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Geary III, William C.
; REGISTRATION NUMBER: 31,359
; REFERENCE/DOCKET NUMBER: UAG-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..845
US-08-589-446-3

Query Match 55.7%; Score 33.4; DB 1; Length 845;
Best Local Similarity 72.9%; Pred. No. 0.017;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 gacatccgcccaggcccccaagaagaccctttcagagactatgtagaccggttctataaaaac 59
||||| | || || || ||||| ||||| | ||||| || ||||| |||| || ||
Db 482 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 540

RESULT 5
US-08-444-882-3
; Sequence 3, Application US/08444882
; Patent No. 5622705
; GENERAL INFORMATION:
; APPLICANT: Morrow, Casey D.
; TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC
; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,882
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Geary III, William C.
; REGISTRATION NUMBER: 31,359
; REFERENCE/DOCKET NUMBER: UAG-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941

```
;   INFORMATION FOR SEQ ID NO:  3:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  845 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;      MOLECULE TYPE:  cDNA
;      FEATURE:
;        NAME/KEY:  CDS
;        LOCATION:  20..845
US-08-444-882-3

  Query Match              55.7%;  Score 33.4;  DB 1;  Length 845;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatccgccaggacccaaggagcccttccgcgactacgtggaccgcttcttcaaagac 59
              ||||   | || || || ||||| |||||  | || || || ||||| |||  || ||
Db        482 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 540

RESULT  6
US-08-389-459A-3
; Sequence 3, Application US/08389459A
; Patent No. 5817512
;  GENERAL INFORMATION:
;     APPLICANT:  Morrow, Casey D. and Porter, Donna, C.
;     TITLE OF INVENTION:  ENCAPSIDATED RECOMBINANT POLIOVIRUS
;     TITLE OF INVENTION:  NUCLEIC ACID AND METHODS OF MAKING AND
;     TITLE OF INVENTION:  USING SAME
;     NUMBER OF SEQUENCES:  23
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  LAHIVE & COCKFIELD
;       STREET:  60 STATE STREET, SUITE 510
;       CITY:  BOSTON
;       STATE:  MASSACHUSETTS
;       COUNTRY:  USA
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/389,459A
;       FILING DATE:  15-FEB-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/087,009
;       FILING DATE:  01-JUL-1993
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Silveri, Jean M.
;       REGISTRATION NUMBER:  39,030
;       REFERENCE/DOCKET NUMBER:  UAB-004CP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 227-7400
;       TELEFAX:  (617) 227-5941
;   INFORMATION FOR SEQ ID NO:  3:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  845 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;      MOLECULE TYPE:  cDNA
;      FEATURE:
;        NAME/KEY:  CDS
;        LOCATION:  20..845
US-08-389-459A-3

  Query Match              55.7%;  Score 33.4;  DB 1;  Length 845;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatccgccaggacccaaggagcccttccgcgactacgtggaccgcttcttcaaagac 59
              ||||   | || || || ||||| |||||  | || || || ||||| |||  || ||
Db        482 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 540

RESULT  7
US-08-987-867A-3
; Sequence 3, Application US/08987867A
; Patent No. 6063384
;  GENERAL INFORMATION:
;     APPLICANT:  C. Morrow et al.
;     TITLE OF INVENTION:  ENCAPSIDATED RECOMBINANT VIRAL
;     TITLE OF INVENTION:  NUCLEIC ACID AND METHODS OF MAKING AND
;     TITLE OF INVENTION:  USING SAME
;     NUMBER OF SEQUENCES:  23
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  LAHIVE & COCKFIELD
;       STREET:  28 STATE STREET
;       CITY:  BOSTON
;       STATE:  MASSACHUSETTS
;       COUNTRY:  USA
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/987,867A
;       FILING DATE:  09-DEC-1997
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/087,009
;       FILING DATE:  01-JUL-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Myers, Louis
;       REGISTRATION NUMBER:  35,965
;       REFERENCE/DOCKET NUMBER:  UAB-004CPDV
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 227-7400
;       TELEFAX:  (617) 742-4214
;   INFORMATION FOR SEQ ID NO:  3:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  845 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;      MOLECULE TYPE:  cDNA
;      FEATURE:
;        NAME/KEY:  CDS
;        LOCATION:  20..845
US-08-987-867A-3

  Query Match              55.7%;  Score 33.4;  DB 3;  Length 845;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatccgccaggacccaaggagcccttccgcgactacgtggaccgcttcttcaaagac 59
              ||||   | || || || ||||| |||||  | || || || ||||| |||  || ||
Db        482 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 540

RESULT  8
5204259-10
;Patent No. 5204259
;    APPLICANT: HELLING,TORSTEN B.;CREVIN, HAKAN,NUNN,
;MICHAEL F.
;    TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
```

```
;     NUMBER OF SEQUENCES: 20
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/344,237
;       FILING DATE: 26-APR-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 191,299
;       FILING DATE: 06-MAY-1988
;       APPLICATION NUMBER:191,229
;       FILING DATE: 06-MAY-1988
;       APPLICATION NUMBER: 206,499
;       FILING DATE: 13-JUN-1988
;       APPLICATION NUMBER: 258,016
;       FILING DATE: 14-OCT-1988
;SEQ ID NO:10:
;       LENGTH:871
5204259-10


  Query Match             55.7%;  Score 33.4;  DB 6;  Length 871;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 qacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqcttcttcaaqac 59
           ||||| | || || || ||||| |||||  | ||||| || ||||| |||| || ||
Db     472 qacataaqacaaqqaccaaaqqaaccctttagagactatgtagaccggttctataaaac 530


RESULT   9
5204259-8
;Patent No. 5204259
;     APPLICANT: BELTING,TORSTEN B.,CREVIN, HAKAN,NUNN,
;MICHAEL F.
;     TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING HIV
;ANTIGENS
;     NUMBER OF SEQUENCES: 20
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/344,237
;       FILING DATE: 26-APR-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 191,299
;       FILING DATE: 06-MAY-1988
;       APPLICATION NUMBER:191,229
;       FILING DATE: 06-MAY-1988
;       APPLICATION NUMBER: 206,499
;       FILING DATE: 13-JUN-1988
;       APPLICATION NUMBER: 258,016
;       FILING DATE: 14-OCT-1988
;SEQ ID NO:8:
;       LENGTH:1021
5204259-8


  Query Match             55.7%;  Score 33.4;  DB 6;  Length 1021;
  Best Local Similarity   72.9%;  Pred. No. 0.017;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 qacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqcttcttcaaqac 59
           ||||| | || || || ||||| |||||  | ||||| || ||||| |||| || ||
Db     472 qacataaqacaaqqaccaaaqqaaccctttagagactatgtagaccggttctataaaac 530


RESULT  10
US-08-463-210-5
; Sequence 5, Application US/08463210
; Patent No. 6001977
;  GENERAL INFORMATION:
;     APPLICANT:  CHANG, Nancy T.
;     APPLICANT:  GALLO, Robert C.
;     APPLICANT:  WONG STAAL, Flossie
;     TITLE OF INVENTION:  CLONING AND EXPRESSION OF HTLV III DNA
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Morgan & Finnegan, L.L.P.
;       STREET:  345 Park Avenue
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10154-0053
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/463,210
;       FILING DATE:  05-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 06/693,866
;       FILING DATE:  23-JAN-1985
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 06/659,339
;       FILING DATE:  10-OCT-1984
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Serunian, Leslie A.
;       REGISTRATION NUMBER:  35,353
;       REFERENCE/DOCKET NUMBER:  2029-4192/RDS
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5362 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  HTLV-III
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  1..5362
;       OTHER INFORMATION:  /standard_name= "Clone BH5"
;       OTHER INFORMATION:  /note= "Corresponds to nucleotide positions 222 to
;       OTHER INFORMATION:  5585 in Figure 3 of US 06/693,866 (parent)"
US-08-463-210-5


  Query Match             55.7%;  Score 33.4;  DB 4;  Length 5362;
  Best Local Similarity   72.9%;  Pred. No. 0.02;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 qacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqcttcttcaaqac 59
           ||||| | || || || ||||| |||||  | ||||| || ||||| |||| || ||
Db     960 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1018


RESULT  11
US-08-850-049-128
; Sequence 128, Application US/08850049
; Patent No. 5965726
;  GENERAL INFORMATION:
;     APPLICANT:
;     APPLICANT:
;     APPLICANT:
;     TITLE OF INVENTION:  METHOD OF ELIMINATING
;     TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;     NUMBER OF SEQUENCES:  130
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/850,049
;      FILING DATE:  02-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/050,478
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  128:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7228 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
US-08-850-049-128

  Query Match             55.7%;  Score 33.4;  DB 2;  Length 7228;
  Best Local Similarity   72.9%;  Pred. No. 0.02;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatccgcgaggacccaaggaacccttccgcgactacgtagaccgcttcttcaagac 59
              |||||  | |  || || ||||| ||||   | |||||  || ||||| ||||  || ||
Db       1578 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1636

RESULT 12
US-08-850-049-129
; Sequence 129, Application US/08850049
; Patent No. 5965726
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/850,049
;      FILING DATE:  02-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/050,478
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  129:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7228 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
US-08-850-049-129

  Query Match             55.7%;  Score 33.4;  DB 2;  Length 7228;
  Best Local Similarity   72.9%;  Pred. No. 0.02;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 gacatccgcgaggacccaaggaacccttccgcgactacgtagaccgcttcttcaagac 59
              |||||  | |  || || ||||| ||||   | |||||  || ||||| ||||  || ||
Db       1578 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1636

RESULT 13
US-08-050-478-128
; Sequence 128, Application US/08050478
; Patent No. 5972596
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/050,478
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  MORRY, MARY J.
;       REGISTRATION NUMBER:  34,398
;       REFERENCE/DOCKET NUMBER:  2026-4006US1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)758 4800
;       TELEFAX:  (212)751-6849
;   INFORMATION FOR SEQ ID NO:  128.
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7228 BASE PAIRS
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
US-08-050-478-128


  Query Match             55.7%;  Score 33.4;  DB 2;  Length 7228;
  Best Local Similarity   72.9%;  Pred. No. 0.02;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gacatccagccaggcccccaaggaqcccttcccgcgactacgtqgaccgcttcttcaaqac 59
           |||||  | || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db    1578 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1636


RESULT  14
US-08-050-478-129
; Sequence 129, Application US/08050478
; Patent No. 5972596
;  GENERAL INFORMATION:
;     APPLICANT:
;     APPLICANT:
;     APPLICANT:
;     TITLE OF INVENTION:  METHOD OF ELIMINATING
;     TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;     NUMBER OF SEQUENCES:  130
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/050,478
;       FILING DATE:  26-OCT-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/02908
;       FILING DATE:  29-MAR-1993
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/858,747
;       FILING DATE:  27-MAR-1992
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  MORRY, MARY J.
;       REGISTRATION NUMBER:  34,398
;       REFERENCE/DOCKET NUMBER:  2026-4006US1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)758-4800
;       TELEFAX:  (212)751-6849
;   INFORMATION FOR SEQ ID NO:  129:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7228 BASE PAIRS
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
US-08-050-478-129


  Query Match             55.7%;  Score 33.4;  DB 2;  Length 7228;
  Best Local Similarity   72.9%;  Pred. No. 0.02;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gacatccagccaggcccccaaggaqcccttcccgcgactacgtqgaccgcttcttcaaqac 59
           |||||  | || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db    1578 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1636


RESULT  15
US-09-414-117-128
; Sequence 128, Application US/09414117
; Patent No. 6291664
;  GENERAL INFORMATION:
;     APPLICANT:
;     APPLICANT:
;     APPLICANT:
;     TITLE OF INVENTION:  METHOD OF ELIMINATING
;     TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;     NUMBER OF SEQUENCES:  130
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/414,117
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/850,049
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/02908
;       FILING DATE:  29-MAR-1993
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/858,747
;       FILING DATE:  27-MAR-1992
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  MORRY, MARY J.
;       REGISTRATION NUMBER:  34,398
;       REFERENCE/DOCKET NUMBER:  2026-4006US1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)758-4800
;       TELEFAX:  (212)751-6849
;   INFORMATION FOR SEQ ID NO:  128:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7228 BASE PAIRS
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
US 09 414 117-128


  Query Match             55.7%;  Score 33.4;  DB 4;  Length 7228;
  Best Local Similarity   72.9%;  Pred. No. 0.02;
  Matches   43;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gacatccagccaggcccccaaggaqcccttcccgcgactacgtqgaccgcttcttcaaqac 59
           |||||  | || || || ||||| |||||  | ||||| || ||||| ||||  || ||
Db    1578 GACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAC 1636
```

Search completed: July 2, 2002, 21:59:00
Job time: 4027 secs

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 23:09:45 ; Search time 291.1 Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-09-475-704-2
Perfect score: 60
Sequence: 1 gacatccgccaggcccaa..........tggaccgcttcttcaagacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: [illegible]
2: [illegible]
3: [illegible]
4: [illegible]
5: [illegible]
6: [illegible]
7: [illegible]
8: [illegible] NA1987.DAT:*
9: [illegible]
10: [illegible]
11: [illegible]
12: [illegible]
13: [illegible]
14: [illegible]
15: [illegible]
16: [illegible]
17: [illegible]
18: [illegible]
19: [illegible]
20: [illegible]
21: [illegible]
22: [illegible]
23: [illegible]
24: [illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	60	100.0	60	21	AAA51608	HIV Gag major homo
2	60	100.0	1509	21	AAA51610	HIV synthetic Gag
3	60	100.0	1509	21	AAA51626	HIV codon-optimize
4	60	100.0	1608	22	AAD09487	Human immunodefici
5	60	100.0	1914	22	AAD09488	Human immunodefici
6	60	100.0	2493	22	AAD09489	Human immunodefici
7	60	100.0	4350	22	AAD09490	Human immunodefici
8	58.4	97.3	60	21	AAA70426	HIV p55 Gag Major
9	58.4	97.3	1268	21	AAA70417	HIV Gag common reg
10	58.4	97.3	1503	21	AAZ52051	Codon optimised HI
11	58.4	97.3	1515	21	AAA70412	Synthetic HIV Gag
12	58.4	97.3	1853	21	AAA70413	HIV Gag protease c
13	58.4	97.3	1865	21	AAA70475	HIV Gag protease c
14	58.4	97.3	1865	21	AAA70476	HIV Gag protease c
15	58.4	97.3	2031	21	AAA70415	Synthetic HIV Gag/
16	58.4	97.3	4319	21	AAA70414	HIV Gag polymerase
17	58.4	97.3	4472	21	AAA70472	HIV bicistronic ex
18	58.4	97.3	4608	21	AAA70473	HIV bicistronic ex
19	58.4	97.3	4665	21	AAA70471	HIV bicistronic ex
20	58.4	97.3	4766	21	AAA70470	HIV bicistronic ex
21	58.4	97.3	8908	21	AAZ52059	Packaging constru
22	55.2	92.0	60	21	AAA51607	HIV Gag major homo
23	55.2	92.0	1479	21	AAA51609	HIV synthetic Gag
24	55.2	92.0	1479	21	AAA51625	HIV codon optimize
25	47.8	79.7	4307	20	AAZ08740	HIV gagpol SYNgp c
26	47.8	79.7	4307	21	AAA93972	Human immunodefici
27	47.8	79.7	4307	22	AAH43696	Codon optimised HI
28	47.8	79.7	4307	22	AAC86876	Nucleotide sequenc
29	47.8	79.7	4327	21	AAA93984	HIV partial leader
30	47.8	79.7	4353	21	AAA93983	HIV partial leader
31	47.8	79.7	4642	21	AAA93982	HIV complete leade
32	47.8	79.7	9772	22	AAH43699	Codon optimised ga
33	47.2	78.7	1482	22	AAF31286	Humanised HIV 1 ga
34	47.2	78.7	1532	19	AAV45173	Synthetic HIV gag
35	47.2	78.7	1532	22	AAF31283	Humanised HIV 1 ga
36	47.2	78.7	4288	22	AAH20868	HIV 1 subtype (C/B
37	45.6	76.0	1539	24	AAI67899	Nucleotide sequenc
38	45.6	76.0	1539	24	AAI67900	Nucleotide sequenc
39	45.6	76.0	5981	24	AAI67898	Nucleotide sequenc
40	45.6	76.0	[illegible]	24	[illegible]	Nucleotide sequenc
41	[illegible]	[illegible]	[illegible]	24	[illegible]	Nucleotide sequenc
42	36.2	[illegible]	[illegible]	17	AAT14[illegible]	Attenuated HIV 1 s
43	[illegible]	[illegible]	[illegible]	21	AAF57929	HIV 1 non subtype
44	[illegible]	[illegible]	[illegible]	21	AAF57926	HIV 1 non subtype
45	34.4	[illegible]	[illegible]	21	[illegible]	Major homology reg

ALIGNMENTS

RESULT 1
AAA51608
ID AAA51608 standard; DNA; 60 BP.
XX
AC AAA51608;
XX
DT 31-OCT-2000 (first entry)
XX
DE HIV Gag major homology region nucleotides 841-900.
XX
KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA immunisation; [illegible] producing cell line; antigen presentation; ss.
XX
OS Human immunodeficiency virus.
XX
PN WO200039304-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31273.
XX
PR 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barnett S, Zur Megede J;
XX
DR WPI; 2000-452461/39.
XX
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal

```
PT   especially human against HIV
XX
PS   Claim 1; Page 92; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic
CC   type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC   useful in DNA immunization, generation of packaging cell lines and
CC   production of Gag and/or Env containing proteins. Synthetic Env and Gag
CC   expression cassettes exhibit increased potency for induction of
CC   cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC   self-assemble into non-infectious virus-like particles which are used as
CC   a matrix for the proper presentation of an antigen entrapped or
CC   associated to the immune system of the host.
XX
SQ   Sequence 60 BP; 12 A; 24 C; 15 G; 9 T; 0 other;


  Query Match             100.0%;  Score 60;  DB 21;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 2e-09;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaggqccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gacatccgccaggqccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60


RESULT   2
AAA51610
ID   AAA51610 standard; DNA; 1509 BP.
XX
AC   AAA51610;
XX
DT   31-OCT-2000  (first entry)
XX
DE   HIV synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus type C strain AF110967
OS   Synthetic.
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31273.
XX
PR   31 DEC 1998;   98US-0114495.
PR   01-SEP-1999;   99US-0152195.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT   Env polypeptide and the polypeptide useful for immunizing a mammal
PI   especially human against HIV
XX
PS   Claim 2; Page 93; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic
CC   type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC   useful in DNA immunization, generation of packaging cell lines and
CC   production of Gag and/or Env containing proteins. Synthetic Env and Gag
CC   expression cassettes exhibit increased potency for induction of
CC   cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC   self assemble into non-infectious virus-like particles which are used as
CC   a matrix for the proper presentation of an antigen entrapped or
CC   associated to the immune system of the host.
XX
SQ   Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 other;


  Query Match             100.0%;  Score 60;  DB 21;  Length 1509;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-09;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaggqccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gacatccgccaggqccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 900


RESULT   3
AAA51626
ID   AAA51626 standard; DNA; 1509 BP.
XX
AC   AAA51626;
XX
DT   31-OCT-2000  (first entry)
XX
DE   HIV codon optimized synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus type C strain AF110967.
OS   Synthetic.
XX
PN   WO200039304-A2.
XX
PD   06 JUL 2000.
XX
PF   30-DEC-1999;   99WO-US31273.
XX
PR   31 DEC-1998;   98US-0114495.
PR   01-SEP-1999;   99US-0152195.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT   Env polypeptide and the polypeptide useful for immunizing a mammal
PI   especially human against HIV
XX
PS   Disclosure; Page 104; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic
CC   type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC   useful in DNA immunization, generation of packaging cell lines and
CC   production of Gag and/or Env containing proteins. Synthetic Env and Gag
CC   expression cassettes exhibit increased potency for induction of
CC   cytotoxic T lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC   self assemble into non-infectious virus-like particles which are used as
CC   a matrix for the proper presentation of an antigen entrapped or
CC   associated to the immune system of the host.
XX
SQ   Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 other;


  Query Match             100.0%;  Score 60;  DB 21;  Length 1509;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-09;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaggqccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gacatccgccaggqccccaaggagccctccgcgactacgtggaccgcttcttcaagacc 900


RESULT   4
AAD09487
```

```
ID   AAD09487 standard; DNA; 1608 BP.
XX
AC   AAD09487;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus A (HIV A) gene.
XX
KW   Human immunodeficiency virus A; HIV A; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector;
KW   MVA; cytotoxic T lymphocyte; CTL; epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             19..1602
FT                   /*tag=  a
FT                   /product= "HIV A immunogen"
XX
PN   WO200147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB04984.
XX
PR   23-DEC-1999;   99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (IIAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T,  Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAE04825.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T lymphocyte epitopes of HIV
PT   protein
XX
PS   Claim 29; Fig 2A; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject.
CC   The invention also relates to method of stimulating anti-HIV immune
CC   response in a human subject which comprises administering one or more
CC   times an amount of nucleic acid molecule sufficient to prime an immune
CC   response to the immunogen, or else may be packaged within a delivery
CC   means, such as a modified vaccinia virus Ankara (MVA) to boost the immune
CC   response to common portion of the immunogens. The present DNA sequence
CC   encodes human immunodeficiency virus A immunogen (HIV A) fusion protein
CC   construct. HIV A immunogen consists of about 73% of gag protein fused to
CC   a string of 25 partially overlapping human CTL epitopes. The gag domain
CC   of HIV A contains p24 and p17 in an order reversed to the viral gag
CC   p17-p24-p15 polyprotein.
XX
SQ   Sequence 1608 BP; 314 A; 598 C; 458 G; 238 T; 0 other;

  Query Match             100.0%;  Score 60;  DB 22;  Length 1608;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-09;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 gacatcggcagggcccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 [illegible]

RESULT 5
AAD09488
ID   AAD09488 standard; DNA; 1914 BP.
XX
AC   AAD09488;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus IA (HIV IA) gene.
XX
KW   Human immunodeficiency virus IA; HIV IA; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector;
KW   MVA; cytotoxic T lymphocyte; CTL; epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             13..1911
FT                   /*tag=  a
FT                   /product= "HIV IA immunogen"
XX
PN   WO200147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB04984.
XX
PR   23-DEC-1999;   99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (IIAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T,  Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAE04826.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T lymphocyte epitopes of HIV
PT   protein -
XX
PS   Example 2; Fig 6B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject.
CC   The invention also relates to method of stimulating anti-HIV immune
CC   response in a human subject which comprises administering one or more
CC   times an amount of nucleic acid molecule sufficient to prime an immune
CC   response to the immunogen, or else may be packaged within a delivery
CC   means, such as a modified vaccinia virus Ankara (MVA) to boost the immune
CC   response to common portion of the immunogens. The present DNA sequence
CC   encodes human immunodeficiency virus IA immunogen (HIV IA) fusion
CC   protein construct. HIV IA construct is based on HIV A immunogen.
CC   The HIV A immunogen consists of about 73% of gag protein fused to
CC   a string of 25 partially overlapping human CTL epitopes. The gag domain
CC   of HIV A contains p24 and p17 in an order reversed to the viral gag
CC   p17-p24-p15 polyprotein. HIV IA shares the same design rationale with
CC   HIV A construct, but additionally includes the HIV-1 clade A tat sequence
CC   positioned between the gag protein and synthetic human CTL epitopes.
XX
SQ   Sequence 1914 BP; 383 A; 728 C; 534 G; 269 T; 0 other;
```

```
  Query Match             100.0%;  Score 60;  DB 22;  Length 1914;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-09;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaqqgcccccaaqqaqccctccqcqactacqtggaccqcttcttcaaqacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 gacatccgccaqqgcccccaaqqaqccctccqcqactacqtggaccqcttcttcaaqacc 525


RESULT   6
AAD09489
ID   AAD09489 standard; DNA; 2493 BP.
XX
AC   AAD09489;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus AeT (HIV AeT) gene.
XX
KW   Human immunodeficiency virus AeT; HIV AeT; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector;
KW   MVA; cytotoxic T lymphocyte; CTL; epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             2179..2493
FT                   /*tag= a
FT                   /product= "HIV AeT immunogen"
XX
PN   WO200147955 A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB04984
XX
PR   23-DEC-1999;   99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (IIAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T,  Mcmichael AJ;
XX
DR   WPI; 2001 418221/44.
DR   P-PSDB; AAE04827.
XX
PI   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
PT   protein  -
XX
PS   Example 2; Fig 7B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation, and a synthetic polypeptide comprising
CC   human cytotoxic T lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject.
CC   The invention also relates to method of stimulating anti-HIV immune
CC   response in a human subject which comprises administering one or more
CC   times an amount of nucleic acid molecule sufficient to prime an immune
CC   response to the immunogen, or else may be packaged within a delivery
CC   means, such as a modified vaccinia virus Ankara (MVA) to boost the immune
CC   response to common portion of the immunogen. The present DNA sequence
CC   encodes human immunodeficiency virus AeT immunogen (HIV AeT) fusion
CC   protein construct. HIV AeT construct is based on HIV A immunogen.
CC   The HIV A immunogen consists of about 73% of gag protein fused to
CC   a string of 25 partially overlapping human CTL epitopes. The gag domain
CC   of HIV A contains p24 and p17 in an order reversed to the viral gag
CC   p17-p24-p15 polyprotein. HIV AeT shares the same design rationale with
CC   HIV A construct, but additionally includes the HIV 1 clade A tat sequence
CC   which is present on the same construct, but expressed as a separate
CC   polypeptide, by virtue of the inclusion of an internal ribosome entry
CC   site.
XX
SQ   Sequence 2493 BP; 516 A; 887 C; 685 G; 405 T; 0 other;


  Query Match             100.0%;  Score 60;  DB 22;  Length 2494;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-09;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaqqgcccccaaqqaqccctccqcqactacqtggaccqcttcttcaaqacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 gacatccgccaqqgcccccaaqqaqccctccqcqactacqtggaccqcttcttcaaqacc 525


RESULT   7
AAD09490
ID   AAD09490 standard; DNA; 4450 BP.
XX
AC   AAD09490;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus PPA immunogen.
XX
KW   Human immunodeficiency virus; PPA; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector;
KW   MVA; cytotoxic T-lymphocyte; CTL; epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             13..4450
FT                   /*tag= a
FT                   /product= "PPA immunogen"
XX
PN   WO200147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB04984.
XX
PR   23-DEC-1999;   99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (IIAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T,  Mcmichael AJ;
XX
DR   WPI; 2001 418221/44.
DR   P-PSDB; AAE04828.
XX
PI   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
PT   protein  -
XX
PS   Claim 40; Fig 8B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation, and a synthetic polypeptide comprising
CC   human cytotoxic T lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
```

```
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject.
CC   The invention also relates to method of stimulating anti HIV immune
CC   response in a human subject which comprises administering one or more
CC   times an amount of nucleic acid molecule sufficient to prime an immune
CC   response to the immunogen, or else may be packaged within a delivery
CC   means, such as a modified vaccinia virus Ankara (MVA) to boost the immune
CC   response to common portion of the immunogens. The present DNA sequence
CC   encodes human immunodeficiency virus PPA immunogen fusion protein
CC   construct. PPA construct is based on HIV A immunogen. The HIV A immunogen
CC   consists of about 73% of gag protein based to a string of 25 partially
CC   overlapped human CTL epitopes. The gag domain of HIV A contains p24 and
CC   p17 in an order reversed to the viral gag p17-p24 p15 polyprotein. PPA
CC   shares the same design rationale with HIV A construct, but additionally
CC   includes the HIV 1 clade A tat sequence, C pol protein, N pol protein,
CC   rev and nef proteins.
XX
SQ   Sequence 4350 BP; 888 A; 1630 C; 1259 G; 573 T; 0 other;


  Query Match             100.0%;  Score 60;  DB 22;  Length 4350;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-09;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaggcccccaggaggccttccgccgactacgtgaccgcttctttccaggacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 gacatccgccaggcccccaggaggccttccgccgactacgtgaccgcttctttccaggacc 525


RESULT   8
AAA70426
ID   AAA70426 standard; DNA; 60 BP.
XX
AC   AAA70426;
XX
DT   28-NOV-2000  (first entry)
XX
DE   HIV p55 Gag Major Homology Region, codon sequence.
XX
KW   HIV 1; AIDS; Gag; vaccine, expression cassette; ss.
XX
OS   Human immunodeficiency virus type 1.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..60
FT                   /*gen=  a
FT                   /partial
FT                   /product= "HIV p55 Gag Major Homology Region protein"
FT                   /note= "No stop codon given"
XX
PN   WO200039302-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;  99WO-US31245.
XX
PR   31-DEC-1998;  98US-0114495.
PR   01-DEC-1999;  99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J, Srivastava I, Lian Y, Barnoy K, Liu H;
PI   Greer C, Selby M, Walker C;
XX
DR   WPI; 2000-452400/39.
DR   P-PSDB; AAR14214.
XX
PT   Expression cassettes encoding the human immunodeficiency virus (HIV)
PT   Gag containing polypeptide useful for vaccinating against HIV
PT   infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS   Claim 1; Page 44; 91pp; English.
XX
CC   The present invention relates to synthetic HIV Gag expression cassettes.
CC   The Gag protein of HIV is needed for the assembly of virus like
CC   particles. In addition, the Gag protein is involved in many stages of the
CC   HIV life cycle, including assembly, virion maturation after particle
CC   release and early post entry steps in viral replication. The expression
CC   cassette may be used for the recombinant expression of HIV
CC   Gag-polypeptides which may then be used to vaccinate against HIV
CC   infection and acquired immunodeficiency syndrome (AIDS). The present
CC   sequence is the coding sequence of HIV p55 Gag Major Homology Region.
CC   This sequence is located within the p24 CA sequence of Gag. Mutations in
CC   this region can severely impair particle production.
XX
SQ   Sequence 60 BP; 13 A; 24 C; 15 G; 8 T; 0 other;


  Query Match             97.4%;  Score 58.4;  DB 21;  Length 60;
  Best Local Similarity   98.4%;  Pred. No. 6e-09;
  Matches   59;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gacatccgccaggcccccaggaggccttccgccgactacgtgaccgcttctttccaggacc 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gacatccgccaggcccccaggaggccttccgccgactacgtgaccgcttctttccaggacc 60


RESULT   9
AAA70417
ID   AAA70417 standard; DNA; 1268 BP.
XX
AC   AAA70417;
XX
DT   28-NOV-2000  (first entry)
XX
DE   HIV Gag common region codon sequence.
XX
KW   HIV 1; AIDS; Gag; vaccine, expression cassette; ss.
XX
OS   Human immunodeficiency virus type 1
XX
PN   WO200039302-A2
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;  99WO-US31245.
XX
PR   31-DEC-1998;  98US-0114495.
PR   01-DEC-1999;  99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J, Srivastava I, Lian Y, Barnoy K, Liu H;
PI   Greer C, Selby M, Walker C;
XX
DR   WPI; 2000-452400/39.
XX
PT   Expression cassettes encoding the human immunodeficiency virus (HIV)
PT   Gag containing polypeptide useful for vaccinating against HIV
PT   infections and acquired immunodeficiency syndrome (AIDS)
XX
PS   Claim 2; Page 343; 591pp; English.
XX
CC   The present invention relates to synthetic HIV Gag expression cassettes.
CC   The Gag protein of HIV is needed for the assembly of virus like
CC   particles. In addition, the Gag protein is involved in many stages of the
CC   HIV life cycle, including assembly, virion maturation after particle
CC   release and early post entry steps in viral replication. The expression
CC   cassette may be used for the recombinant expression of HIV
CC   Gag polypeptides which may then be used to vaccinate against HIV
CC   infection and acquired immunodeficiency syndrome (AIDS). The present
CC   sequence is a common region found in Gag codon sequences; Gag protease
CC   and Gag-polymerase.
XX
SQ   Sequence 1268 BP; 274 A; 449 C; 411 G; 135 T; 0 other;
```

```
  Query Match             97.3%;  Score 58.4;  DB 21;  Length 1268;
  Best Local Similarity   98.3%;  Pred. No. 7.4e-09;
  Matches   59;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqctttctacaaqacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     862 gacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqctttctacaaqacc 921


RESULT  10
AAZ52051
ID   AAZ52051 standard; DNA; 1503 BP.
XX
AC   AAZ52051;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Codon optimised Human immunodeficiency virus gag coding region.
XX
KW   HIV; gag; packaging cell line; lentivirus; retroviral vector particle;
KW   gal; gene therapy; gene replacement; vaccine; biochemical reagent;
KW   codon optimisation; ds.
XX
OS   Human immunodeficiency virus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1503
FT                   /*tag= a
FT                   /product= "gag protein"
XX
PN   WO200015819-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20675.
XX
PR   11-SEP-1998;   98US-0100022.
PR   10-SEP-1998;   98US-0100063.
XX
PA   (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI   Gray JT,  Mulligan RC;
XX
DR   WPI, 2000-271455/23.
DR   P-PSDB, AAY70599.
XX
PT   New packaging cell line for producing a viral accessory protein
PT   independent HIV derived retroviral vector particles, useful in gene
PT   therapy or gene replacement
XX
PS   Claim 41; Fig 8; 62pp; English.
XX
CC   The patent discloses new packaging cell line for producing a viral
CC   accessory protein independent lentivirus, preferably
CC   human immunodeficiency virus (HIV), derived retroviral vector particles.
CC   The packaging cell line comprises a mammalian cell, a retroviral DNA
CC   comprising a coding sequence for a lentivirus, preferably HIV, gagpol,
CC   where the coding sequence has been mutagenised to improve expression of
CC   the viral gagpol proteins, a second retroviral nucleotide sequence
CC   comprising the coding sequence for a heterologous envelope protein and a
CC   third retroviral nucleotide sequence comprising a DNA sequence of
CC   interest and lentivirus, preferably HIV, cis-acting sequences required
CC   for packaging, reverse transcription and integration.
CC   The packaging cell lines and viral particles can be used for gene
CC   therapy or gene replacement with improved safety. They can also be used
CC   in the development and production of vaccines and biochemical reagents.
CC   The present sequence is a gag coding region of codon optimised HIV
CC   gagpol sequence. This sequence is used in the packaging
CC   cell line. Codon optimisation results in improved expression of the
CC   gagpol protein and reduces the risk of recombination between the transfer
CC   vector and gagpol mRNA.
XX
SQ   Sequence 1503 BP; 329 A; 540 C; 434 G; 201 T; 0 other;


  Query Match             97.3%;  Score 58.4;  DB 21;  Length 1503;
  Best Local Similarity   98.3%;  Pred. No. 7.5e-09;
  Matches   59;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqctttctacaaqacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     850 gacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqctttctacaaqacc 909


RESULT  11
AAA70412
ID   AAA70412 standard; DNA; 1515 BP.
XX
AC   AAA70412;
XX
DT   28-NOV-2000  (first entry)
XX
DE   Synthetic HIV Gag expression cassette coding sequence Gag.ModSE2.
XX
KW   HIV 1; AIDS; Gag; vaccine, expression cassette, ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
PN   WO200039302 A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31245.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-DEC-1999;   99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Ge M,  Mojedo I,  Srivastava I,  Lian Y,  Hartog K,  Liu H;
PI   Greer C,  Selby M,  Walker C;
XX
DR   WPI; 2000-452400/39.
XX
PT   Expression cassettes encoding the human immunodeficiency virus (HIV)
PT   Gag containing polypeptide useful for vaccinating against HIV
PT   infections and acquired immunodeficiency syndrome (AIDS)
XX
PS   Claim 9; Fig 7; 491pp; English.
XX
CC   The present sequence is the coding sequence of a HIV Gag expression
CC   cassette, Gag.ModSE2. The Gag protein of HIV is needed for the assembly
CC   of virus-like particles. In addition, the Gag protein is involved in
CC   many stages of the HIV life cycle, including assembly, virion maturation
CC   after particle release and early post-entry steps in viral replication.
CC   The expression cassette may be used for the recombinant expression of
CC   HIV Gag polypeptides which may then be used to vaccinate against HIV
CC   infection and acquired immunodeficiency syndrome (AIDS).
XX
SQ   Sequence 1515 BP; 329 A; 547 C; 480 G; 159 T; 0 other;


  Query Match             97.3%;  Score 58.4;  DB 21;  Length 1515;
  Best Local Similarity   98.3%;  Pred. No. 7.5e-09;
  Matches   59;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqctttctacaaqacc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     862 gacatccqccaqqqccccaaqqaqccctccqcqactacqtqqaccqctttctacaaqacc 921


RESULT  12
AAA70413
```

```
ID   AAA70413 standard; DNA; 1853 BP.
XX
AC   AAA70413;
XX
DT   28-NOV-2000  (first entry)
XX
DE   HIV Gag protease expression cassette coding sequence GagProt.ModS.
XX
KW   HIV-1; AIDS; Gag protease; vaccine; expression cassette, ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
PN   WO200039302-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31245.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-DEC-1999;   99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Zur Megede J,  Srivastava I,  Lian Y,  Hartog K,  Liu H;
PI   Greer C,  Selby M,  Walker C;
XX
DR   WPI; 2000-452400/39.
XX
PT   Expression cassettes encoding the human immunodeficiency virus (HIV)
PT   Gag-containing polypeptide useful for vaccinating against HIV
PT   infections and acquired immunodeficiency syndrome (AIDS)
XX
PS   Claim 5; Fig 7; 391pp; English.
XX
CC   The present sequence is the coding sequence of a HIV Gag-protease
CC   expression cassette, GagProt.ModS. The Gag protein of HIV is needed for
CC   the assembly of virus-like particles. In addition, the Gag protein is
CC   involved in many stages of the HIV life cycle, including assembly, virion
CC   maturation after particle release and early post entry steps in viral
CC   replication. The expression cassette may be used for the recombinant
CC   expression of HIV Gag-polypeptides which may then be used to vaccinate
CC   against HIV infection and acquired immunodeficiency syndrome (AIDS).
XX
SQ   Sequence 1853 BP; 421 A; 624 C; 580 G; 228 T; 0 other;

  Query Match             97.3%;  Score 58.4;  DB 21;  Length 1853;
  Best Local Similarity   98.3%;  Pred. No. 7.6e-09;
  Matches   59;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 gacatccgccagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        862 gacatccgccagggccccaaggagcccttccgcgactacgtggaccgcttctacaagacc 921

RESULT 13
AAA70475
ID   AAA70475 standard; DNA; 1865 BP.
XX
AC   AAA70475;
XX
DT   28-NOV-2000  (first entry)
XX
DE   HIV Gag protease expression cassette coding sequence GagProtMod.SF2(GP1).
XX
KW   HIV 1; AIDS; Gag protease; vaccine; expression cassette, ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
PN   WO200039302-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31245.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-DEC-1999;   99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Zur Megede J,  Srivastava I,  Lian Y,  Hartog K,  Liu H;
PI   Greer C,  Selby M,  Walker C;
XX
DR   WPI; 2000-452400/39.
XX
PT   Expression cassettes encoding the human immunodeficiency virus (HIV)
PT   Gag containing polypeptide useful for vaccinating against HIV
PT   infections and acquired immunodeficiency syndrome (AIDS)
XX
PS   Claim 5; Fig 69; 391pp; English.
XX
CC   The present sequence is the coding sequence of a HIV Gag-protease
CC   expression cassette, GagProtMod.SF2 (GP1). The Gag protein of HIV is
CC   needed for the assembly of virus like particles. In addition, the Gag
CC   protein is involved in many stages of the HIV life cycle, including
CC   assembly, virion maturation after particle release and early post entry
CC   steps in viral replication. The expression cassette may be used for the
CC   recombinant expression of HIV Gag-polypeptides which may then be used to
CC   vaccinate against HIV infection and acquired immunodeficiency syndrome
CC   (AIDS).
XX
SQ   Sequence 1865 BP; 424 A; 627 C; 583 G; 231 T; 0 other;

  Query Match             97.3%;  Score 58.4;  DB 21;  Length 1865;
  Best Local Similarity   98.3%;  Pred. No. 7.6e-09;
  Matches   59;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 gacatccgccagggccccaaggagcccttccgcgactacgtggaccgcttcttcaagacc 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        868 gacatccgccagggccccaaggagcccttccgcgactacgtggaccgcttctacaagacc 927

RESULT 14
AAA70476
ID   AAA70476 standard; DNA; 1865 BP.
XX
AC   AAA70476;
XX
DT   28-NOV-2000  (first entry)
XX
DE   HIV Gag protease expression cassette coding sequence GagProtMod.SF2(GP2).
XX
KW   HIV 1; AIDS; Gag protease; vaccine; expression cassette; ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
PN   WO200039302-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;   99WO-US31245.
XX
PR   31-DEC-1998;   98US-0114495.
PR   01-DEC-1999;   99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Zur Megede J,  Srivastava I,  Lian Y,  Hartog K,  Liu H;
PI   Greer C,  Selby M,  Walker C;
XX
DR   WPI; 2000-452400/39.
XX
```

PT Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS Claim 5; Fig 70; 391pp; English.
XX
CC The present sequence is the coding sequence of a HIV Gag-protease
CC expression cassette, GagProtMod.SF2(GP2). The Gag protein of HIV is
CC needed for the assembly of virus like particles. In addition, the Gag
CC protein is involved in many stages of the HIV life cycle, including
CC assembly, virion maturation after particle release and early post-entry
CC steps in viral replication. The expression cassette may be used for the
CC recombinant expression of HIV Gag-polypeptides which may then be used to
CC vaccinate against HIV infection and acquired immunodeficiency syndrome
CC (AIDS).
XX
SQ Sequence 1865 BP; 460 A; 583 C; 569 G; 253 T; 0 other;

Query Match 97.3%; Score 58.4; DB 21; Length 1865;
Best Local Similarity 98.3%; Pred. No. 7.6e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gacatccgccaggqcccaaqqaqccctccgcgactacgtqqaccgcttcttcaagacc 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db 868 gacatccgccaggqcccaaqqaqccctccgcgactacgtqqaccgcttctacaagacc 927

RESULT 15
AAA70415
ID AAA70415 standard; DNA; 2031 BP.
XX
AC AAA70415;
XX
DT 28-NOV-2000 (first entry)
XX
DE Synthetic HIV Gag/HCV core fusion coding sequence.
XX
KW HIV-1; AIDS; Gag; vaccine; expression cassette; ss.
XX
OS Chimeric - Human immunodeficiency virus type 1.
OS Chimeric - Hepatitis C virus.
XX
PN WO200039302-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31245.
XX
PR 31-DEC-1998; 98US-0114495.
PR 01-DEC-1999; 99US-0168471.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
PI Greer C, Selby M, Walker C;
XX
DR WPI, 2000-452400/39.
XX
PT Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS Example 1; Pages 341-342; 391pp; English.
XX
CC The present sequence is a HIV Gag/Hepatitis C virus (HCV) core fusion
CC coding sequence. The Gag protein of HIV is needed for the assembly of
CC virus like particles. In addition, the Gag protein is involved in many
CC stages of the HIV life cycle, including assembly, virion maturation after
CC particle release and early post-entry steps in viral replication. The
CC present invention relates to synthetic HIV Gag expression cassettes. The
CC present sequence was cloned and used to generate the expression cassettes
CC of the present invention. The expression cassettes may be used for the
CC recombinant expression of HIV Gag-polypeptides which may then be used to
CC vaccinate against HIV infection and acquired immunodeficiency syndrome
CC (AIDS).
XX
SQ Sequence 2031 BP; 421 A; 707 C; 646 G; 257 T; 0 other;

Query Match 97.4%; Score 58.4; DB 21; Length 2031;
Best Local Similarity 98.4%; Pred. No. 7.6e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gacatccgccaggqcccaaqqaqccctccgcgactacgtqqaccgcttcttcaagacc 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db 862 gacatccgccaggqcccaaqqaqccctccgcgactacgtqqaccgcttctacaagacc 921

Search completed: July 2, 2002, 24:09:46
Job time: 8498 sec